



Structural and functional characterisation of LdcB
LD-carboxypeptidases from
Streptococcus pneumoniae* and *Bacillus subtilis

Christopher Neil Hoyland

Newcastle University
Faculty of Medical Sciences
Institute for Cell and Molecular Biosciences
September 2014

Abstract

The bacterial cell wall surrounds the cytoplasmic membrane and protects the cell against osmolysis in addition to providing shape. The cell wall is comprised of peptidoglycan, repeating units of N-acetyl glucosamine and N-acetyl muramic acid form glycan strands and are crosslinked by short peptides that contain both L- and D-amino acids. Owing to the unique nature of peptidoglycan, and its absence in eukaryotic organisms, the cell wall has become an important target for many antibiotics, including the β -lactams and glycopeptides.

Newly synthesised peptidoglycan contains pentapeptides, which extend from the lactyl moiety of the MurNAc sugar. These chains consist of L-alanine-D- γ -glutamate/glutamine-L-lysine/meso-diaminopimelic acid-D-alanine-D-alanine. The terminal D-alanine is often lost during cell wall maturation, either as a result of the crosslinking reaction, in which the penultimate D-alanine is attached to the side-chain of a neighbouring L-lysine or meso-diaminopimelic acid by an isopeptide bond, or as a consequence of the activities of DD-carboxypeptidases, and results in a tetrapeptide. The tetrapeptide can then be trimmed further to form a tripeptide by the action of LD-carboxypeptidases. Although many DD-carboxypeptidases have been well characterised, the majority of LD-carboxypeptidases that have been studied are active only against peptidoglycan fragments and so cannot be responsible for producing the tripeptides found in the cell wall. Of the LD-carboxypeptidases active against the mature cell wall, DacB (*Streptococcus pneumoniae*), Csd6 (*Helicobacter pylori*) and Pgp2 (*Campylobacter jejuni*), each has been shown to be essential in maintaining cell morphology. It should be noted, however, that neither Csd6 nor Pgp2 share any sequence similarity with DacB and belong to different peptidase families.

This thesis concerns the structural and biochemical characterisation of DacB, herein renamed to LdcB (LD-carboxypeptidase B). The crystal structures of the *apo* form of LdcB from both *S. pneumoniae* and *Bacillus subtilis* were solved, revealing a single domain, globular protein with 2 sub-domains forming a V-shaped cleft in which the active site is located. LdcB binds one zinc ion per monomer, located at the bottom of the active site, and is a member of the LAS (lysostaphin, D-Ala-D-Ala peptidases, sonic hedgehog) family of metalloproteins. Additionally, the activity of LdcB as an LD-carboxypeptidase was confirmed and the crystal structure of LdcB from *S. pneumoniae*

was solved in complex with a product mimic, M-Tri-Lys(D-Asn), revealing the molecular basis for peptidoglycan recognition in this family of enzymes. Finally, the affinity of LdcB for zinc and copper has been determined and it has been shown that catalysis is not inhibited by the substitution of zinc by copper or cobalt.

Contents

1 Introduction	1
1.1 The bacterial cell envelope	2
1.1.1 Overview	2
1.1.2 The cell envelope as a drug target	3
1.2 Gram-positive cell wall structure	4
1.2.1 Peptidoglycan	4
1.2.2 Teichoic acids	6
1.3 Peptidoglycan synthesis	8
1.4 Peptidoglycan incorporation	9
1.4.1 Glycosyltransfer	11
1.4.2 Transpeptidation	12
1.5 Localisation of the HMW PBPs	14
1.5.1 Localisation patterns	14
1.5.2 Determinants of localisation	19
1.6 Transpeptidation substrate availability is controlled by carboxypeptidases	20
1.7 Other carboxypeptidase functions	24
1.7.1 Peptidoglycan recycling	24
1.7.2 Antibiotic resistance	28
1.8 LdcB/DacB	33
1.9 Aims of this study	35
2 Materials and methods	36
2.1 Molecular biology	37
2.1.1 Chemical suppliers	37
2.1.2 General buffers and solutions	37
2.1.3 Growth media	37
2.1.4 Strains and plasmids	37
2.1.5 Preparation of competent cells	39
2.1.6 Transformation of competent cells	39
2.1.7 QuikChange mutagenesis of <i>SpLdcB</i>	39
2.1.8 Visualisation of mutagenic products by agarose gel electrophoresis	40
2.1.9 Purification of plasmids from small-scale cultures	41
2.1.10 DNA sequencing	41

2.1.11 Analysis of proteins by SDS-PAGE electrophoresis.....	41
2.2 Protein overexpression and purification	42
2.2.1 Large-scale overexpression and purification of target proteins (native and Se-Met substituted)	42
2.2.2 Purification of target proteins by metal affinity chromatography (Ni ⁺ -NTA).....	42
2.2.3 Removal of hexahistidine tag by thrombin/TEV cleavage.....	43
2.2.4 Purification of proteins by size exclusion chromatography (Superdex 75).....	43
2.2.5 Extraction of native metals by dialysis against EDTA and subsequent buffer exchange.....	44
2.3 Biophysical characterisation of purified proteins	45
2.3.1 Determination of protein concentration.....	45
2.3.2 Circular dichroism.....	46
2.4 Metal analysis of LdcB enzymes.....	47
2.4.1 Inductively coupled plasma mass spectrometry (ICP-MS)	47
2.4.2 Calculating binding affinities for d-block metals and <i>BsLdcB</i>	47
2.5 Biochemical characterisation of LdcB proteins.....	48
2.5.1 Carboxypeptidase activity	48
2.5.1.1 Thin-layer chromatography (TLC).....	48
2.5.1.2 Quantitative release of D-alanine.....	49
2.5.2 Identification of natively bound ligands	51
2.5.2.1 Electrospray mass spectrometry	51
2.5.2.2 Marfey's reagent and reverse-phase HPLC	51
2.6 Crystallisation of target proteins	53
2.6.1 Crystallisation trials	53
2.6.2 In-house screening of crystals for diffraction.....	53
2.6.3 Data collection.....	54
2.6.3.1 Data collection for molecular replacement.....	54
2.6.3.2 SAD data collection.....	54
2.6.3.3 Data collection for anomalous difference maps	54
2.6.4 Data processing	54
2.6.5 Determination of an initial model by molecular replacement.....	56
2.6.6 Heavy atom substructure solution (SAD)	57
2.6.7 Calculation of anomalous difference maps	58
2.6.8 Model building and refinement.....	58
2.6.9 Model validation.....	59

3 Structural characterisation of LdcB from <i>S. pneumoniae</i> and <i>B. subtilis</i>	60
3.1 Introduction	61
3.2 Crystallisation trials	61
3.2.1 Crystallisation trials of native <i>Sp</i> LdcB and <i>Bs</i> LdcB	61
3.3 Crystallisation of Se-Met substituted <i>Sp</i> LdcB and the addition of zinc	62
3.3.1 Incorporation of selenomethionine	62
3.3.2 Addition of zinc and effect on crystallisation of LdcB	63
3.3.3 Data processing and structure solution of Se-Met substituted <i>Sp</i> LdcB	67
3.3.4 Data processing and molecular replacement of <i>Bs</i> LdcB	71
3.4 The structure of LdcB	75
3.4.1 The structures of <i>Sp</i> LdcB and <i>Bs</i> LdcB	75
3.4.2 Identification of native-bound ligands	78
3.4.3 Comparing the LdcB structures	83
3.4.4 Comparison of <i>Sp</i> LdcB and <i>Bs</i> LdcB with LAS (lysostaphin, D-Ala-D-Ala metallopeptidases, sonic hedgehog) family proteins	86
3.4.5 Comparing <i>Sp</i> LdcB and <i>Bs</i> LdcB with vancomycin resistance enzymes	89
3.4.6 Comparing LdcB enzymes to other LD-carboxypeptidases	92
3.5 Summary	93
 4 The structure of <i>Sp</i>LdcB complexed with MurNAc-Ala-D-γ-Gln-Lys(D-Asn) & the biochemical characterisation of LdcB enzymes	 95
4.1 Introduction	96
4.2 Co-crystallisation	96
4.2.1 Selection of ligands for co-crystallisation studies	96
4.2.2 Crystallisation trials of <i>Sp</i> LdcB with a range of products and substrates	97
4.2.3 The structure of <i>Sp</i> LdcB with M-Tri-Lys(D-Asn) bound to the active site	101
4.2.4 Modelling of the tetrapeptide substrate	114
4.3 Biochemical and biophysical assays on the native enzyme	116
4.3.1 Thin-layer chromatography (TLC) of <i>Sp</i> LdcB reaction products	116
4.3.2 Construction and characterisation of active site mutants	121
4.3.2.1 Confirmation of mutant folding	121
4.3.2.2 Thin-layer chromatography (TLC) of mutant <i>Sp</i> LdcB reaction products	122
4.4 Kinetic characterisation of <i>Sp</i> LdcB for Ala-D-γ-Gln-Lys-D-Ala	125
4.5 A Proposed catalytic mechanism	127
4.6 Summary	129

5 Metal binding in LdcB enzymes	132
5.1 Introduction	133
5.2 Loading non-native metals.....	133
5.3 The affinity of LdcB for zinc, cobalt and copper	135
5.4 Crystal structures of the <i>Sp</i> LdcB:metal complexes.....	141
5.4.1 Crystallisation and structure solution.....	141
5.4.2 Crystal structures of the <i>Sp</i> LdcB:cobalt complex	146
5.4.3 Crystal structures of the <i>Sp</i> LdcB:copper complex.....	148
5.5 The effects of d-block metals on catalysis	149
5.6 Summary	153
 6 Summary and Discussion	 155
6.1 Summary of the major findings of this thesis	156
6.2 Other LdcB structures	156
6.3 The LdcB represents an ancestral form of the vancomycin resistance enzymes	160
6.4 Convergent evolution between peptidases.....	164
6.5 The physiological role of LdcB.....	166
6.6 Concluding remarks and future work	168
 References	 170
 Appendix	 189
Appendix 1: General reagents, buffers and growth media	190
Appendix 2: Plasmid map of pET28a	192
Appendix 3: Primers used in this study	193
Appendix 4: Co-crystallisation conditions of <i>Sp</i> LdcB with a range of ligands	194
 Papers arising from post-graduate studies.....	 195

List of figures

Figure		Page
Figure 1.1	Schematic representation of the Gram-negative (left) and Gram-positive (right) cell envelopes.	2
Figure 1.2	Chemical schematic of the cell wall stem pentapeptide.	5
Figure 1.3	Schematic representation of the crosslinks found within the bacterial cell wall.	6
Figure 1.4	The chemical structures of <i>B. subtilis</i> and <i>S. aureus</i> wall teichoic acid.	7
Figure 1.5	Schematic representation of peptidoglycan biosynthesis.	9
Figure 1.6	The crystal structures of a class A PBP (top) vs. a class B PBP (bottom).	10
Figure 1.7	Schematic representation of peptidoglycan incorporation.	11
Figure 1.8	Chemical schematic of the cell wall crosslinks.	14
Figure 1.9	Schematic representation of the growth and division methods employed by rod-shape, spherical cocci and ovococci cells.	16
Figure 1.10	Schematic showing the cleavage points of various cell wall carboxypeptidases.	21
Figure 1.11	The crystal structure of a DacA (PBP3) from <i>S. pneumoniae</i> , a class C PBP.	21
Figure 1.12	Electron microscopy showing the morphologies of wild-type <i>S. pneumoniae</i> (A) and a $\Delta dacA$ strain (B), taken from Schuster <i>et al.</i> , (1990).	22
Figure 1.13	Scanning electron microscopy pictures of N6 <i>H. pylori</i> (A) and $\Delta hdpA$ N6 <i>H. pylori</i> (B), taken from Bonis <i>et al.</i> , (2010).	23
Figure 1.14	Chemical structures of D-Ala-D-Ala (top left) and Lys-D-Ala (top right) and their non-hydrolysable mimics (bottom left and right).	26

Figure 1.15	(A) The structures of <i>NaLdcA</i> and <i>PaLdcA</i> superposed; (B) Surface representation of <i>NaLdcA</i> with catalytically important residues shown.	28
Figure 1.16	Schematic showing the mode of action of vancomycin.	29
Figure 1.17	The superposed crystal structures of VanXYg and VanXYc.	31
Figure 1.18	The active site of VanXYc with a phosphinate mimic of D-Ala-D-Ala bound (A) and D-Ala-D-Ala bound (B).	32
Figure 1.19	Phase-contrast and fluorescent vancomycin images of wild-type and $\Delta ldcB$ encapsulated (IU1690) and unencapsulated (IU1945) <i>S. pneumoniae</i> , taken from Barendt <i>et al.</i> , (2011).	33
Figure 1.20	The localisation patter of 3xFLAG tagged LdcB, taken from Barendt <i>et al.</i> , (2011).	34
Figure 2.1	Representative mutagenesis agarose gel.	40
Figure 2.2	Representative chromatogram of <i>SpLdcB</i> elution from a Ni-NTA column.	43
Figure 2.3	Representative chromatogram of <i>SpLdcB</i> elution from a Superdex S75 HiLoad 16/60 size exclusion column.	44
Figure 2.4	Representative purification gel.	44
Figure 2.5	BSA concentration standard curve.	45
Figure 2.6	Reaction scheme for the conversion of D-alanine into lactate.	50
Figure 2.7	Marfey's reagent derivatisation of L and D alanine.	52
Figure 3.1	Initial crystal forms of <i>SpLdcB</i> (A) and <i>BsLdcB</i> (B).	62
Figure 3.2	Metal content analysis of purified <i>SpLdcB</i> .	64
Figure 3.3	Metal content analysis of <i>SpLdcB</i> following zinc addition.	64
Figure 3.4	Improved crystal form of <i>SpLdcB</i> .	65

Figure 3.5	Metal content analysis (MCA) fluorescence scan of <i>Sp</i> LdcB crystals.	66
Figure 3.6	An example diffraction pattern produced by Se-Met labelled <i>Sp</i> LdcB crystals.	67
Figure 3.7	The output of SHELXC.	68
Figure 3.8	The output of SHELXD.	69
Figure 3.9	The output of SHELXE.	69
Figure 3.10	Ramachandran plot of the Se-Met <i>Sp</i> LdcB structure.	71
Figure 3.11	<i>Bs</i> LdcB plate-like crystals.	72
Figure 3.12	An example diffraction pattern produced by <i>Bs</i> LdcB.	72
Figure 3.13	Ramachandran plot of the <i>Bs</i> LdcB structure.	74
Figure 3.14	Structure of LdcB.	75
Figure 3.15	Structure of the LdcB active site.	76
Figure 3.16	Conservation of zinc-coordinating residues.	77
Figure 3.17	Sequence alignment of a range of LdcB sequences with VanY, VanXYg and VanXYc	77
Figure 3.18	Unknown density in the LdcB active site.	78
Figure 3.19	Electrospray mass spectra of <i>Sp</i> LdcB native (top) and denatured (bottom).	80
Figure 3.20	Reverse-phase HPLC chromatograms for FDAA-linked L-alanine (A), D-alanine (B) and unknown ligand (C) and FDAA by itself (D).	81
Figure 3.21	D- and L-Alanine fitting into the unknown <i>Sp</i> LdcB density.	82
Figure 3.22	Density of a phosphate ion in the <i>Bs</i> LdcB active site.	83
Figure 3.23	Comparing the three LdcB structures.	84
Figure 3.24	Zinc forms crystal contacts between <i>Sp</i> LdcB monomers.	86
Figure 3.25	A comparison of the active sites of LdcB and other LAS family	87

members.

Figure 3.26	A comparison of the active sites of <i>SpLdcB</i> and VanXYg.	90
Figure 3.27	<i>SpLdcB</i> superposed onto VanXYg.	91
Figure 3.28	The structure of <i>SpLdcB</i> (A) compared to LdcA from <i>P. aeruginosa</i> (B) and <i>N. aromaticivorans</i> (C).	93
Figure 4.1	Crystals of <i>SpLdcB</i> co-crystallised with (A) Tri-DAP; (B) M-Tri-Lys(D-Asn); (C) M-Tri-Lys.	98
Figure 4.2	Ramachandran plot of the <i>SpLdcB</i> -ligand structure.	100
Figure 4.3	Schematic representations of the chemical structures of the true LdcB reaction product (A) and the product mimic (B).	102
Figure 4.4	Cartoon representation of ligand-bound <i>SpLdcB</i>	103
Figure 4.5	Interactions between the M-Tri-Lys(D-Asn) substrate and <i>SpLdcB</i> at subsite S ₃ .	104
Figure 4.6	Interactions between the M-Tri-Lys(D-Asn) substrate and <i>SpLdcB</i> at subsite S ₂ .	105
Figure 4.7	The structure of MurNAc-tetrapeptide (A) compared to a mimic in which D-γ-Gln is substituted by D-Gln (B).	106
Figure 4.8	Interactions between the M-Tri-Lys(D-Asn) substrate and <i>SpLdcB</i> at subsite S ₁ .	107
Figure 4.9	Interactions between the M-Tri-Lys(D-Asn) substrate and <i>SpLdcB</i> at subsite S ₁ '.	107
Figure 4.10	Schematic summary of the interactions made between <i>SpLdcB</i> and the product mimic, M-Tri-Lys(D-Asn).	108
Figure 4.11	Sequence alignment showing residue conservation in a range of LdcB and VanY-like enzymes.	109
Figure 4.12	A comparison of the ligand-bound active sites carboxypeptidases.	110
Figure 4.13	Proposed path of cross-linked cell wall peptides.	111

Figure 4.14	L-Lysine (A) vs. meso- diaminopimelic acid (B) containing stem peptides.	112
Figure 4.15	Conformational changes in <i>SpLdcB</i> on ligand binding.	114
Figure 4.16	B-factor analysis of <i>SpLdcB</i> in the closed (A) and open (B) conformations.	115
Figure 4.17	The modelled path of the tetrapeptide substrate of <i>SpLdcB</i> .	115
Figure 4.18	Chemical structures of substrates used in this study.	116
Figure 4.19	Activity of <i>SpLdcB</i> against peptidoglycan tetrapeptides.	117
Figure 4.20	HPLC chromatograms of mucopeptides from <i>S. pneumoniae</i> strains, obtained after incubating PG with or without <i>SpLdcB</i> in the presence of Zn^{2+} or EDTA, followed by digestion with cellosyl and reduction with sodium borohydride.	118
Figure 4.21	Activity of LdcB homologues against peptidoglycan pentapeptide and D-Ala-D-Ala dipeptide.	119
Figure 4.22	The predicted interactions of MurNAc-pentapeptide with <i>SpLdcB</i> .	120
Figure 4.23	Circular dichroism (CD) spectra of <i>SpLdcB</i> and three mutants.	122
Figure 4.24	Activity of wild-type (WT) <i>SpLdcB</i> compared to three active site mutants.	122
Figure 4.25	Time-course TLC of WT, E204A and E204H <i>SpLdcB</i> .	123
Figure 4.26	<i>In silico</i> mutagenesis of <i>SpLdcB</i> E204A and E204H.	124
Figure 4.27	Michaelis-Menten curves for <i>BsLdcB</i> (A), <i>SpLdcB</i> (B) and <i>BaLdcB</i> (C).	126
Figure 4.28	The active sites of <i>SpLdcB</i> and pancreatic carboxypeptidase A (PDB ID: 1HDU (Cho <i>et al.</i> , 2002) are compared.	128
Figure 4.29	The proposed catalytic mechanism of LdcB.	130
Figure 5.1	Representative metal content analysis of EDTA treated <i>SpLdcB</i> .	134

Figure 5.2	Representative metal content analysis of apo- <i>SpLdcB</i> following incubation with copper (left) or cobalt (right).	135
Figure 5.3	Representative data used to determine <i>BsLdcB</i> affinity for (A) Zn^{2+} , (B) Cu^{2+} and (C) Co^{2+} .	137
Figure 5.4	ITC isotherm of NMB0315 for zinc taken from Wang <i>et al.</i> (2011)	140
Figure 5.5	Crystal forms of <i>SpLdcB</i> :cobalt (left) and <i>SpLdcB</i> :copper (right).	142
Figure 5.6	An example diffraction pattern of <i>SpLdcB</i> :cobalt (left) and <i>SpLdcB</i> :copper (right).	142
Figure 5.7	Anomalous difference maps.	145
Figure 5.8	Apo <i>SpLdcB</i> with zinc bound superposed onto <i>SpLdcB</i> :cobalt.	146
Figure 5.9	Metal binding in <i>SpLdcB</i> :cobalt superposed onto the apo structure of <i>SpLdcB</i> with zinc bound.	147
Figure 5.10	Chain A (left) and chain B (right) of <i>SpLdcB</i> :copper superposed onto the structure of <i>SpLdcB</i> with M-Tri-Lys(D-Asn) bound.	148
Figure 5.11	The metal binding site of <i>SpLdcB</i> :copper, compared to apo and ligand-bound <i>SpLdcB</i>	149
Figure 5.12	Activity of <i>SpLdcB</i> loaded with zinc, cobalt or copper against peptidoglycan tetrapeptides	150
Figure 5.13	HPLC chromatograms of mucopeptides from <i>S. pneumoniae</i> strains, obtained after incubating PG with or without <i>SpLdcB</i> loaded with zinc, cobalt or copper, followed by digestion with cellosyl and reduction with sodium borohydride.	151
Figure 6.1	Ligand-bound <i>SpLdcB</i> compared to the open conformation of DacB (PDBid: 4D0Y).	157
Figure 6.2	The position of Glu204 in the crystal structures of apo <i>SpLdcB</i> (PDBid: 40X5), ligand-bound <i>SpLdcB</i> (PDBid: 40XD) and the open form of DacB (4D0Y).	158
Figure 6.3	The position of Trp206 in open and closed DacB.	159

Figure 6.4	The position of the carbonyl oxygen of L-Lys in a model of L-Lys-D-Ala bound to DacB, taken from Abdullah <i>et al.</i> , (2014).	159
Figure 6.5	The hypothesised bond between Trp206 and the tetrapeptide is not possible in the ligand-bound enzyme conformation	160
Figure 6.6	Phylogram comparing LdcB sequences with VanY and VanXYg from <i>Enterococcus faecium</i> , and VanXYc from <i>Enterococcus gallinarum</i> .	161
Figure 6.7	The pentapeptide model projecting above the surface of <i>SpLdcB</i> .	163
Figure 6.8	The active site of CPA (A) and <i>SpLdcB</i> (B)	165
Appendix	Plasmid map of pET28a	192
Figure 2.1		

List of tables

Table		Page
Table 2.1	List of <i>E. coli</i> strains used in this study.	38
Table 2.2	Plasmids and constructs used in this study.	38
Table 2.3	Quickchange reaction conditions.	40
Table 2.4	A list of all structures solved by molecular replacement, the models used and their sequence similarity.	57
Table 3.1	Summary of X-ray data collection statistics for initial <i>SpLdcB</i> crystals.	62
Table 3.2	Summary of X-ray data collection statistics for <i>SpLdcB</i> .	68
Table 3.3	Summary of X-ray data refinement statistics for <i>SpLdcB</i> .	70
Table 3.4	Summary of X-ray data collection statistics for <i>BsLdcB</i> .	73
Table 3.5	Summary of X-ray data refinement statistics for <i>BsLdcB</i> .	74
Table 3.6	Table summary of a section of DALI results.	87
Table 3.7	RMSD scores of the three LdcB enzymes (top) with VanXYg, VanXYc and VanX (side).	91
Table 4.1	The chemical structures of commercially available cell wall fragments. All fragments were purchased from InvivoGen (http://www.invivogen.com)	97
Table 4.2	Summary of representative X-ray data collection statistics for <i>SpLdcB</i> with ligands.	99
Table 4.3	Summary of X-ray data refinement statistics for <i>SpLdcB</i> + M-Tri-Lys(D-Asn).	101
Table 4.4	RMSD values between Ligand-bound <i>SpLdcB</i> chains (4OXD).	103
Table 5.1	Summary of X-ray data collection statistics for <i>SpLdcB</i> :cobalt above (High) and below (Low) the K-edge of the ion.	144

Table 5.2	Summary of X-ray data collection statistics for <i>SpLdcB</i> :copper above (High) and below (Low) the K-edge of the ion.	144
Table 5.3	Summary of X-ray data refinement statistics for <i>SpLdcB</i> :cobalt and <i>SpLdcB</i> :copper.	145
Appendix Table 3.1	List of mutagenesis primers used in this study. Mutations are underlined, with specific base changes in bold.	193
Appendix Table 4.1	Crystallisation conditions for <i>SpLdcB</i> with a range of ligands.	194

List of abbreviations

Abbreviation	Definition
Abs	Absorbance
ATP	Adenosine triphosphate
<i>Ba</i>	<i>Bacillus anthracis</i>
bp	Base pairs
<i>Bs</i>	<i>Bacillus subtilis</i>
CD	Circular dichroism
CPA	Carboxypeptidase A
CSGID	Center for Structural Genomics of Infectious Diseases
Cx	Carbon at position x, where x represents an integer
Da	Dalton
DAO	D-amino acid oxidase
DacB	D-Ala-D-Ala carboxypeptidase B
DLS	Diamond Light Source
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide
ϵ	Extinction coefficient
EDTA	Ethylenediaminetetraacetic acid
FADH ₂	Flavin adenine dinucleotide
FDAA	1-fluoro-2-4-dinitrophenyl-5-L-alanine amide, a.k.a Marfey's reagent
<i>g</i>	Acceleration due to gravity, 9.81 m/s ²
GFP	Green fluorescence protein
GlcNAc	N-acetylglucosamine
GT/GTase	Glycosyltransfer/glycosyltransferase
h	Hour
HMW	High molecular weight
HPLC	High performance liquid chromatography
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
ICP-MS	Inductively-coupled plasma mass spectrometry
IPTG	Isopropyl β -D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
kb	Kilobase
kDa	Kilodalton
KEGG	Kyoto encyclopedia of genes and genomes

LAS	Lysostaphin, D-Ala-D-Ala peptidases, sonic hedgehog
LB	Luria Bertani broth
LdcA	LD-carboxypeptidase A
LdcB	LD-carboxypeptidase B
LDH	lactate dehydrogenase
LMW	Low molecular weight
LPS	Lipopolysaccharide
ManNAc	N-acetylmannosamine
meso-DAP	Meso-diaminopimelic acid
MGT	Monofunctional glycosyltransferase
MPD	2-Methyl-2,4-pentanediol
MR	Molecular replacement
M-Tri-Lys(D-Asn)	N-acetylmuramic acid-L-alanine-D-γ-glutamine-L-lysine, with D-asparagine linked to the side-chain of the lysine via an isopeptide bond.
MurNAc	N-acetylmuramic acid
NADPH	Nicotinamide adenine dinucleotide phosphate
NADH	Nicotinamide adenine dinucleotide
Ni ⁺ -NTA	Nickel nitrilotriacetic acid
NMR	Nuclear magnetic resonance
OD ₆₀₀	Optical density at 600 nm
PAR	4-(2-pyridylazo)resorcinol
PAGE	Polyacrylamide gel electrophoresis
PBP	Penicillin binding protein
PCR	Polymerase chain reaction
PDB	Protein data bank
PDB id	Protein data bank ascension number
PEG	Polyethylene glycol
PEP	Phosphoenolpyruvate
<i>Pfu</i>	<i>Pyrococcus furiosus</i>
PNK	Polynucleotide kinase
RMSD	Root mean square deviation
rpm	Revolutions per minute
SAD	Single-wavelength anomalous dispersion
SDS	Sodium dodecyl sulphate
Se-Met	Selenomethionine
S-layer	Surface layer

<i>Sp</i>	<i>Streptococcus pneumoniae</i>
Tetra-D	Ala-D-Gln-L-Lys-D-Ala
Tetra-L	Ala-D-Gln-L-Lys-L-Ala
TEV	Tobacco etch virus
TLC	Thin-layer chromatography
TP/TPase	Transpeptidation/transpeptidase
Tris	Tris(hydroxymethyl)aminomethane
UDP	Uridine diphosphate
UV	Ultra-violet
WTA	Wall teichoic acid

Standard amino acid abbreviations were used

Acknowledgements

I would like to thank my supervisor Rick Lewis for his support, firstly through my MRes and then through my PhD. He took a risk in taking on a marine biology graduate with only a minor biochemical background and it is thanks to his advice and training that I have been able to accomplish this PhD. I would also like to extend my thanks to Rob, who has taught me a great deal about working in a structural biology lab in addition to working closely with me on several projects and providing the clones used in this thesis.

I also wish to show my appreciation to other members of SBL, both past and present, who have helped me with a variety of things, from Arnaud's help with data collection and structure solution to Joe's advice on structure solution and the useful discussion provided by Vince, Jon, Lorraine, Tom, Cecilia and Will.

I am grateful to Waldemar Vollmer and Christine Aldridge for their work on this project, assessing the activity of purified LdcB on the cell walls of various bacteria. Likewise, I must extend my thanks to Djalal Meziane-Cherif, who collaborated to provide a quantitative analysis of LdcB activity and provided the clone for the *Bacillus anthracis* LdcB homologue, and Kevin Waldron, without whom the ICP-MS analysis would not have been possible.

I would like to thank my family and friends for all their support during my studies.

“If the road is easy, you're likely going the wrong way.”

- Terry Goodkind

Chapter 1:

Introduction

1.1 The bacterial cell envelope

1.1.1 Overview

Bacteria can be broadly separated into two groups based upon the configuration of their cellular envelope, the Gram-negatives and the Gram-positives. The cellular envelope of Gram-negative bacteria comprises a thin coat of peptidoglycan, one or two layers thick, sandwiched between two membranes, the inner phospholipid bilayer and the outer lipopolysaccharide (LPS) layer (**Figure 1.1**) (Koch, 2003). The outer membrane provides bacteria with resistance to toxic molecules, provides an anchor for extracellular proteins and is the reason that these organisms are classed as Gram-negative, as the outer membrane prevents crystal violet, the Gram stain, from binding to the bacterial cell wall.

Gram-positive bacteria, on the other hand, lack an outer membrane and the protection against external stresses it provides (Koch, 2003). In compensation, Gram-positive bacteria have a peptidoglycan coat that is much thicker than that of the Gram-negatives (**Figure 1.1**) (Shockman and Barrett, 1983). This thick wall is sufficiently strong to withstand the turgor of the cell, which in Gram-positive organisms can be higher than 20 atmospheres (Mitchell and Moyle, 1957), and provides the bacteria with a scaffold onto which additional components may be built. These additional components include various proteins, a capsule/surface layer (S-layer) and anionic polymers, i.e. the wall teichoic acids (WTA) (Schäffer and Messner, 2005).

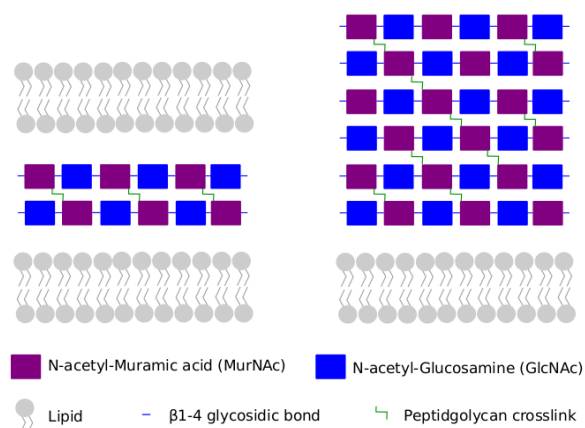


Figure 1.1: Schematic representation of the Gram-negative (left) and Gram-positive (right) cell envelopes. Lipopolysaccharides, membrane associated proteins and anionic polymers are not included in the respective cell envelopes for clarity.

1.1.2 The cell envelope as a drug target

The cell envelopes of Gram-positive and Gram-negative bacteria protect from stresses that would otherwise cause lysis, such as osmotic shock. As such the disruption of the cell envelope has been seen as an ideal method of dealing with bacterial infection, initially by nature and later by man. For example, the first antibiotic used, penicillin, was discovered by Fleming in 1929 as the product of the fungus *Penicillium chrysogenum* (previously *Penicillium notatum*) (Fleming, 1929). Penicillin was shown, through the use of ^{14}C -labelling, to function by preventing the peptide crosslinking of peptidoglycan that is an essential facet of the architecture of the cell wall (Yocum *et al.*, 1980). Since the time of Fleming, penicillin has been adapted artificially to increase its effectiveness; ampicillin, for example, is a derivative that contains an amino group, aiding its penetration of the outer membrane of Gram-negative bacteria. In addition, a number of other drugs with varying modes of action, from the glycopeptides that also target the peptide crosslinks of peptidoglycan, to cationic peptides, such as polymyxin B, that target and disrupt the outer membrane of Gram-negative bacteria, have been introduced as potent antibiotics (Ainsworth *et al.*, 1947; Hirsch *et al.*, 1960; Nagarajan, 1991). Despite this, the overuse of antibiotics in general, and those that target the bacterial cell wall in particular, has resulted in the emergence of antibiotic resistant strains that either possess enzymes that are capable of degrading antibiotics (i.e. β -lactamases), or proteins/components that have a much lower affinity for the commonly used antibiotics (i.e. vancomycin-resistance strains) (Arthur *et al.*, 1996; Paterson and Bonomo, 2005). Furthermore, in the same time period, there has been a decline in the number of new antibiotics reaching the market for several reasons, ranging from inadequate screening techniques (Boucher *et al.*, 2009) to a cynical preference within the pharmaceutical industry to concentrate on other aspects of human health (e.g. expensive cancer treatments, life-long remedies for, for instance, impotence) rather than the rather short-term treatments that often have a low profit margin, such as those required to treat bacterial infection (Hadley, 2004). However, it is clear that the development of new antibiotics is crucial, something that has recently been recognised by not only the UK research councils, which have pledged to form a “war council” to address the issue, and the Chief Medical Officer of the UK, Professor Dame Sally Davies, but by the British public, who have voted for antibiotic resistance to be topic of the Longitudinal prize,

which offers a £10 million reward to whomever can provide a solution to antibiotic resistance. The cell wall remains an attractive target for drug discovery because it avoids the need for any drug to enter and remain in the bacterial cell, and thus if the cell wall is to be targeted, a detailed understanding of its synthesis and maturation is required.

1.2 Gram-positive cell wall structure

1.2.1 Peptidoglycan

Peptidoglycan is often termed the primary cell wall polymer, and it is found in the cell walls of both Gram-positive and Gram-negative bacteria. The three main roles of peptidoglycan are (i) to act as a physical barrier to withstand the turgor of the cell, (ii) to form a molecular mesh to control import and export and (iii) to produce a scaffold to which other components can be attached (Schneewind and Missiakas, 2012). Additional components are required for the assembly of a fully-functional cell wall, for instance the outer membrane of Gram-negative bacteria and the anionic polymers of Gram-positive bacteria (Koch, 2003; Shockman & Barrett, 1983).

The composition of peptidoglycan in Gram-positive bacteria is quite well characterised; Salton and Pavlik (1960) used lysozyme treatment followed by acid hydrolysis and paper chromatography to confirm the presence of N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) in a range of Gram-positive organisms (Salton and Pavlik, 1960). Repeating units of GlcNAc-MurNAc, which are linked by β 1-4 glycosidic bonds, have been shown by atomic force microscopy to form “cables” that run parallel to the cytoplasmic membrane, forming helices around the short axis of the cell (Hayhurst *et al.*, 2008). The digestion of the cell walls of many bacteria by lysozyme, which cleaves the β 1-4 glycosidic bonds in the glycan chains, followed by HPLC separation of the products and mass spectrometry has shown that the glycan chains are crosslinked through short stem peptides present on MurNAc to form a net-like structure (Atrih *et al.*, 1999). In most Gram-positive bacteria, the peptide stem consists of L-Ala-D- γ -Glu/Gln-L-Lys-D-Ala-D-Ala (**Figure 1.2**), with the notable exception of the Bacilli, in which the L-Lysine at position three is replaced by meso-diaminopimelic acid (meso-DAP), which is often amidated (**Figure 1.2**, overleaf) (Vollmer *et al.*, 2008a).

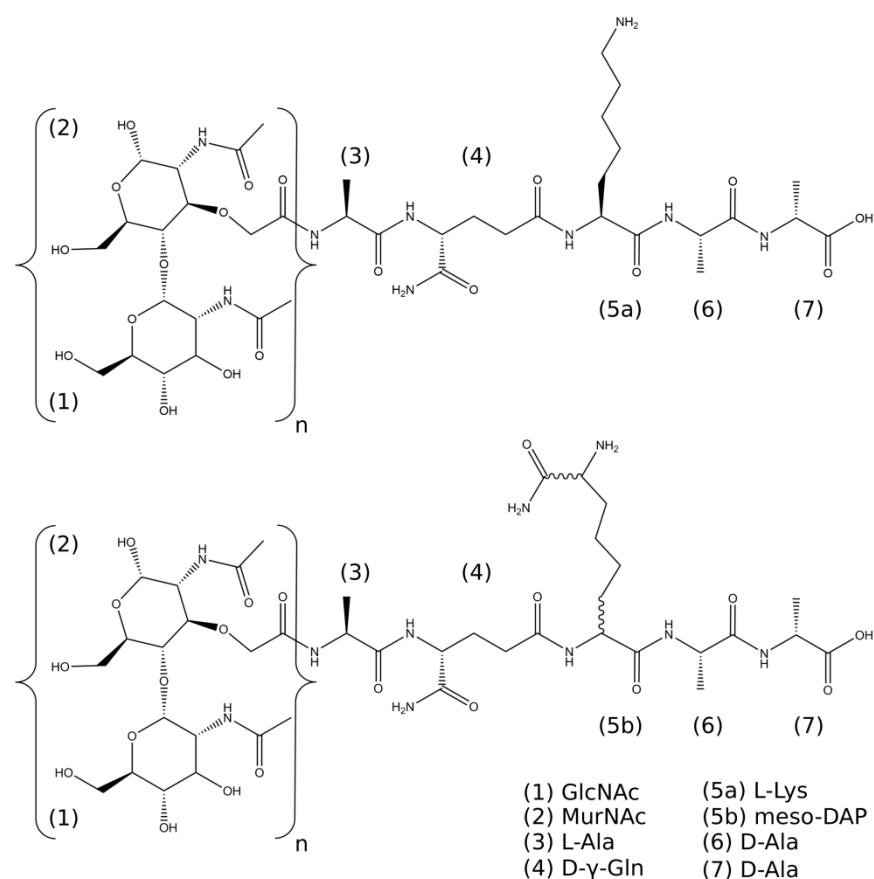


Figure 1.2: Chemical schematic of the cell wall stem pentapeptide. The lysine containing stem pentapeptide can be seen in the top panel, whilst an amidated, meso-DAP containing stem pentapeptide is observed in the lower panel. This figure was produced using ChemBioDraw v14.0.0.117 (PerkinElmer).

In general, crosslinks between peptide stems are formed between the carboxyl group of the residue at position 4 on one strand and the ϵ -amino group on the side-chain of the amino acid at position three on a neighbouring chain, i.e. D-Alanine is linked to L-Lysine (or meso-DAP) in what is referred to as a 3-4 crosslink, though 3-3 and 2-4 crosslinks are also known to occur in some bacteria (**Figure 1.3**, overleaf) (Mainardi *et al.*, 2005; Vollmer *et al.*, 2008a). Furthermore, various modifications can take place once peptidoglycan has been synthesised; the enzymes MurT and GatD have been shown, though the HPLC analysis of cell wall material in $\Delta murT$ and $\Delta gatD$ strains, to catalyse the conversion of D- γ -glutamic acid to D- γ -glutamine in the stem peptide (Figueiredo *et al.*, 2012), the importance of the which is underlined by the observation that genetically modified *Staphylococcus aureus* strains that are unable to amidate the peptide stem show decreased growth and increased susceptibility to both β -lactams and to lysozyme, a major component of the innate immune response (Figueiredo *et al.*, 2012).

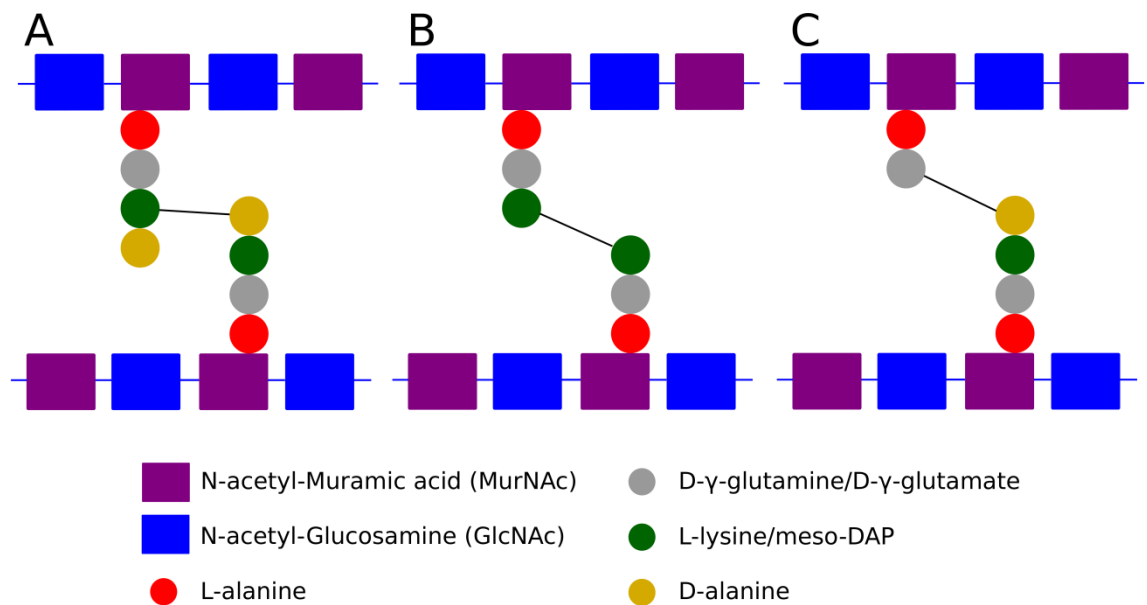


Figure 1.3: Schematic representation of the crosslinks found within the bacterial cell wall. The left-most panel (A) shows the canonical 3-4 crosslink between L-Lys/meso-DAP and D-Ala. The centre panel (B) shows a 3-3 crosslink between D-Ala and a neighbouring D-Ala in two stem tripeptides. The right-most panel (C) shows a 2-4 crosslink between D-γ-Glu/Gln and D-Ala.

1.2.2 Teichoic acids

In addition to peptidoglycan, wall teichoic acids (WTA) comprise roughly 50% of the dry weight of the cell wall of Gram-positive bacteria (Neuhaus and Baddiley, 2003). One proposed function of WTA is to control cation homeostasis, helping to provide a pool of divalent cations for extracellular enzymes and controlling the porosity of the cell wall. Sedimentation studies have demonstrated that, under low ionic strength conditions, WTA adopts a rigid, rod-shaped form, resulting in charge-charge repulsion and increasing the porosity of the wall (Doyle *et al.*, 1974; Neuhaus and Baddiley, 2003). WTA is required for proper biofilm formation in *Staphylococcus aureus*; with bacteria that are unable to synthesise a component of the WTA, the D-alanyl esters, are unable to adhere to artificial surfaces such as glass or plastic, likely as a result of an increase in surface net-negative charge (Gross *et al.*, 2001). Furthermore, WTA is suggested to play a pivotal role in cell growth and division, with *Bacillus subtilis* cells losing their characteristic rod-shape when the enzymes responsible for the final attachment of WTA are depleted (Kawai *et al.*, 2011), indicating that WTA is involved in the regulation of cell elongation in rod-shaped bacteria.

The chemical composition of WTA is known to vary from species to species, with two main WTA types having been identified from early paper chromatography studies in the staphylococci alone; WTA that contains poly-glycerol-phosphate and WTA that contains poly-ribitol-phosphate (Armstrong *et al.*, 1959b). The structure of WTA was elucidated by Armstrong *et al.*, (1959a), whereby the authors treated purified WTA with a variety of chemical compounds that were predicted to variably affect the components of WTA depending on where they were located on the polymer. For example, it was shown that periodate oxidation of WTA from *B. subtilis*, followed by sodium borohydride and acid hydrolysis, resulted in the formation of glycerol, not ethylene glycol as was predicted if glucose was present and attached to the ribitol moiety (Armstrong *et al.*, 1959a). NMR studies have subsequently been used to confirm the structures of WTA predicted by the earlier chemical studies, with the chemical schematics of both poly-glycerol-phosphate and poly-ribitol-phosphates WTA shown in **Figure 1.4**. It has been demonstrated that WTA is attached to the peptidoglycan frame-work through a phosphodiester bond between the C6 hydroxyl of the MurNAc and the anomeric phosphate of the WTA linker unit (Poxton *et al.*, 1978), with the family of enzymes responsible for catalysing this reaction having been identified by Kawai *et al.* (2011).

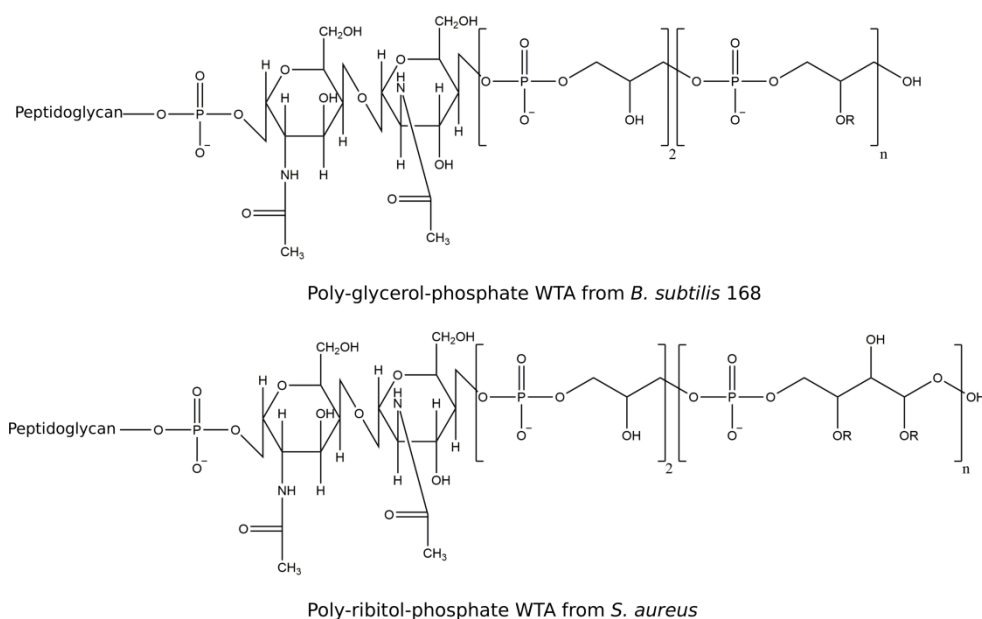


Figure 1.4: The chemical structures of *B. subtilis* and *S. aureus* wall teichoic acid. Both types of WTA contain identical GlcNAc-ManNAc (N-acetyl mannosamine) linker units and have two glycerol phosphate moieties proceeding either a poly-glycerol or poly-ribitol tail. R represents sites of modification, such as the site of D-alanyl esterification. This figure was produced using ChemBioDraw v14.0.0.117 (PerkinElmer)

1.3 Peptidoglycan synthesis

The biosynthesis of peptidoglycan begins in the cytoplasm of the cell with the synthesis of the soluble precursors UDP-GlcNAc and UDP-MurNAc-pentapeptide. UDP-GlcNAc is converted from glucosamine-6-phosphate in a two-step process catalysed by GlmM and GlmU in *Mycobacterium tuberculosis*, or by homologues in other bacteria (Mengin-Lecreulx and van Heijenoort, 1993; Jolly *et al.*, 1999). Enzymes of the Mur family then catalyse the sequential steps required for the conversion of UDP-GlcNAc into UDP-MurNAc-pentapeptide. MurA, first identified by Strominger (1958) and biochemically characterised by Gunetileke and Anwar (1968), has been shown to catalyse the first reaction, transferring enolpyruvate from phosphoenolpyruvate (PEP) to UDP-GlcNAc to form UDP-GlcNAc-enolpyruvate. The next step involves the reduction of UDP-GlcNAc-enolpyruvate to UDP-MurNAc in a NADPH and FADH₂ dependent manner by MurB, as shown by Benson *et al.* (1993) through the use of labelled 4R-[²H]NADPH in NMR experiments. Biochemical assays following the production of radiolabelled UDP-MurNAc-L-[¹⁴C]Ala by recombinant MurC have confirmed that it catalyses the first step of the stem pentapeptide production through the addition of L-alanine to UDP-MurNAc (Liger *et al.*, 1995). Similar biochemical assays, also using radiolabelled substrates have shown that MurD, E and F function to sequentially add the remaining amino acids to the stem pentapeptide, with MurF adding the final D-Ala-D-Ala residues as a dipeptide (Ito and Strominger, 1973; Duncan *et al.*, 1990; Pratviel-Sosa *et al.*, 1991).

The second stage of peptidoglycan synthesis takes place anchored to the cytoplasmic face of the plasma membrane; MraY has been shown to be capable of linking MurNAc-pentapeptide to undecaprenyl pyrophosphate (also known as bactoprenol), through the use of radiolabelled UDP-MurNAc-[¹⁴C]pentapeptide, to form undecaprenyl pyrophosphate-MurNAc-pentapeptide, or lipid I (Umbreit and Strominger, 1972). Analysis of the accumulation of the UDP-GlcNAc and undecaprenyl-pyrophosphate-MurNAc-pentapeptide peptidoglycan precursors in Δ *murG* strains of *E. coli* suggested that MurG catalyses the transfer of GlcNAc onto lipid I to form Lipid II, which was confirmed in the same study through the use of *in vitro* transferase assays (Mengin-Lecreulx *et al.*, 1991). The newly synthesised lipid II is then transferred out of the cell for incorporation into the cell wall, with the flippases

FtsW, MurJ, RodA and SpoVE capable of facilitating the movement of lipid II (Real *et al.*, 2008; Mohammadi *et al.*, 2011; Sieger *et al.*, 2013). Interestingly, MurJ inactive *E. coli* mutants have recently been found to not accumulate the products of lipid II hydrolysis when exposed to ColM, a colicin that degrades lipid II (Sham *et al.*, 2014), suggesting that MurJ acts as the primary lipid II flippase. This contradicts previous *in vitro* results implicating FtsW in this role, with the work by Sham *et al.*, (2014) showing that the inactivation of FtsW does not perturb lipid II export, bringing into question the role of FtsW in lipid II export. The intracellular synthesis and export of peptidoglycan is summarised in **Figure 1.5**.

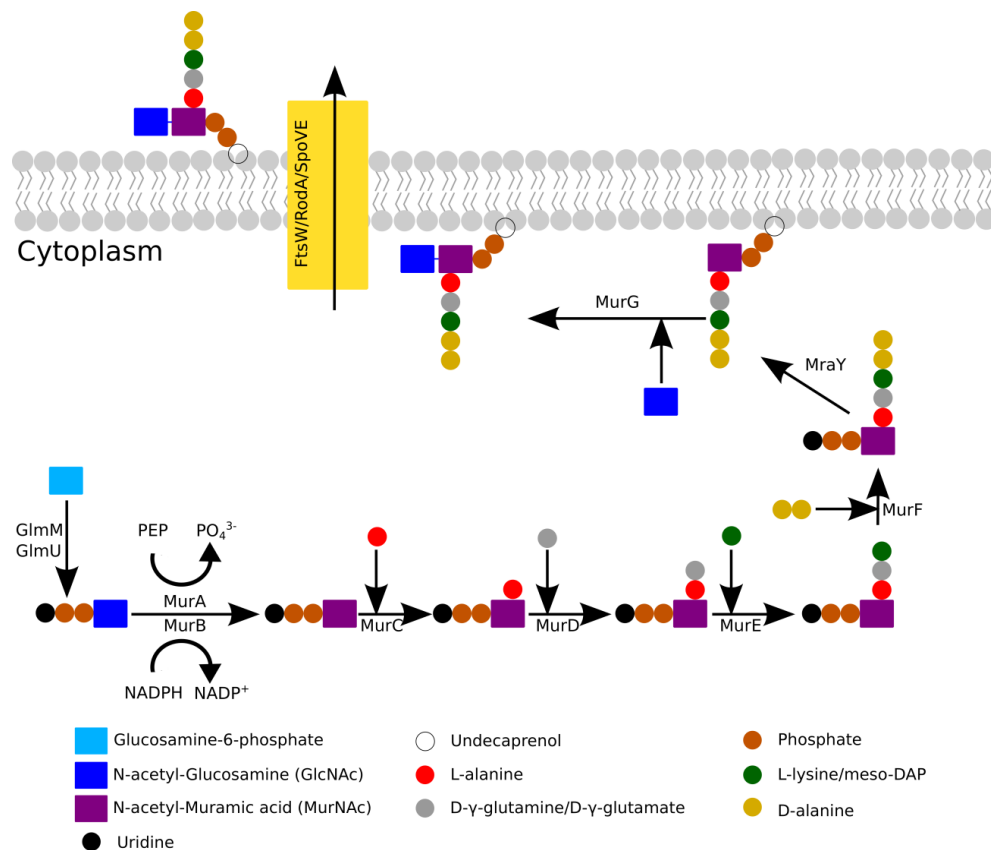


Figure 1.5: Schematic representation of peptidoglycan biosynthesis. The cytoplasmic steps of peptidoglycan synthesis are shown up to export of the Lipid II precursor by FtsW/RodA/SpoVE.

1.4 Peptidoglycan incorporation

Once on the outside of the cellular membrane, the newly synthesised mucopeptide must be incorporated into the existing cell wall mass in a two-step process, glycosyltransfer (GT) followed by transpeptidation (TP). These reactions are mediated

primarily by a group of enzymes known as the high molecular weight (HMW) penicillin binding proteins (PBPs), which can be split into two groups based upon their activity. Class A PBPs possess a GTase domain in addition to their TPase domain, whilst class B PBPs possess only TPase activity, although they do possess an additional domain known as the ‘pedestal’ (**Figure 1.6**). It is thought that the pedestal domain positions the TPase domain away from the membrane and near to the peptidoglycan, although an allosteric site, 60 Å from the active site, has been located on the pedestal domain of PBP2A of *S. aureus* which causes a conformational change and opens the active site of the enzyme when muramic acid or peptidoglycan is bound, suggesting a possible regulatory role for the pedestal domain (Sauvage *et al.*, 2008; Otero *et al.*, 2013). In addition, non-PBP monofunctional glycosyltransferases (MGTs) are known to exist that can perform the glycosyltransferase reaction independently of the PBPs (Di Berardino *et al.*, 1996; Spratt *et al.*, 1996).

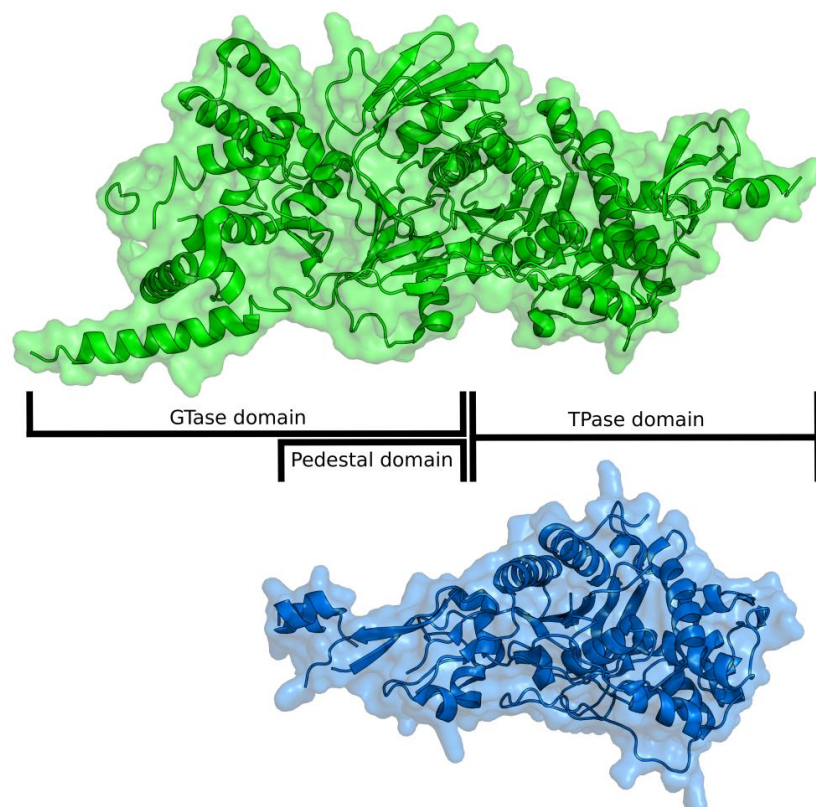


Figure 1.6: The crystal structures of a class A PBP (top) vs. a class B PBP (bottom). The different domains of the PBPs are indicated; both classes possess a TPase domain. Class A PBPs also possess a GTase domain, whilst class B PBPs contain a ‘pedestal’ domain. The class A PBP is represented by PBP1b from *E. coli* (PDBid: 3VMA) (Sung *et al.*, 2009) whilst the class B PBP is PBP3, also from *E. coli* (PDBid: 4BJP) (Sauvage *et al.*, 2014). The PBPs are shown as a cartoon ribbon with a transparent surface. The core RMSD between the TPase domains of the class A and class B PBPs is ~ 1.6 Å.

1.4.1 Glycosyltransfer

The GT reaction inserts the new chain into the wall through the formation of a β 1-4 glycosidic bond between the C1 of a MurNAc sugar in the growing peptidoglycan chain and the 4-OH of a GlcNAc sugar in lipid II (**Figure 1.7**) (for a review on the structural details behind the GTase step see Lovering *et al.* (2008)). The reaction detaches the undecaprenyl pyrophosphate lipid from the MurNAc of the growing peptidoglycan chain, with the liberated undecaprenyl pyrophosphate being converted to undecaprenyl phosphate by UppP before being flipped back across the cytoplasmic membrane, by an as-of-yet unknown process, which may be catalysed by a protein or be facilitated by the lipid-nature of undecaprenyl phosphate, for use in further rounds of peptidoglycan, WTA or other cell envelope component biosynthesis (Tatar *et al.*, 2007).

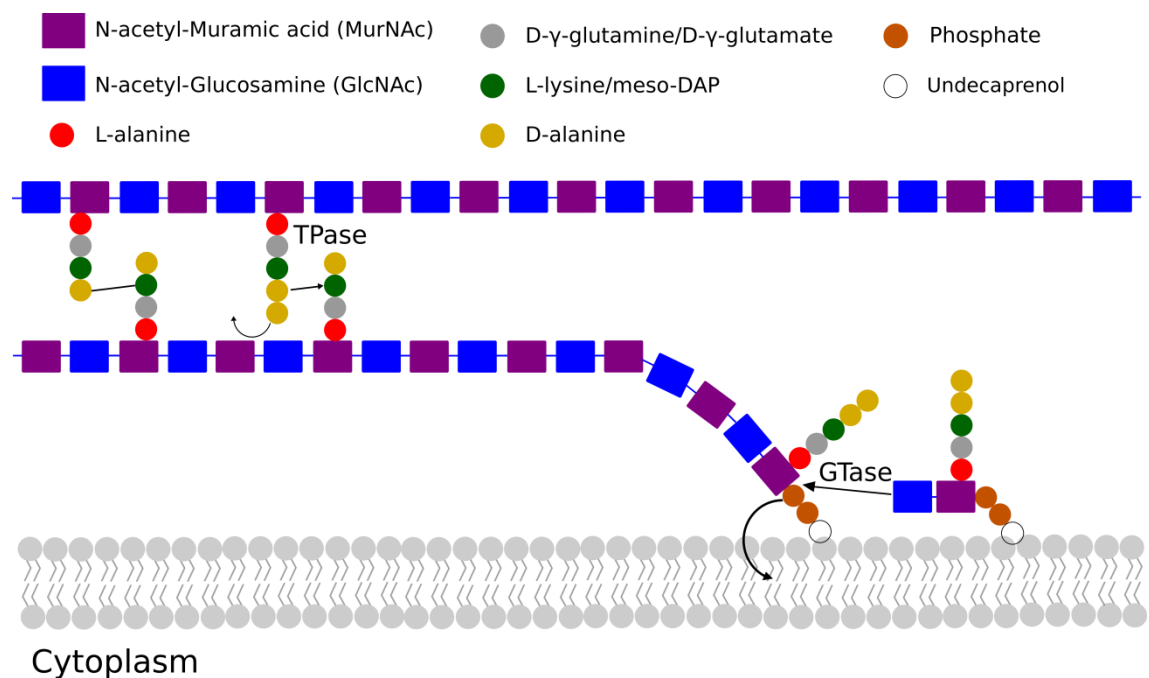


Figure 1.7: Schematic representation of peptidoglycan incorporation. Once on the external leaflet of the cytoplasmic membrane, the PBP enzymes perform the GTase reaction to link the new peptidoglycan monomer with the existing chain. This detaches the undecaprenyl pyrophosphate from the previously added monomer, allowing it to be recycled for further round of synthesis. The PBPs then link adjacent glycan strands together via the TPase reaction, here shown to form a 3-4 crosslink, linking a donor pentapeptide (which becomes a tetrapeptide as a consequence of the reaction) to an accepting tetrapeptide.

1.4.2 Transpeptidation

Once peptidoglycan is incorporated into sugar backbone of the cell wall by the GT reaction, crosslinks are formed between stem peptides of adjacent cell wall muropeptides (Sauvage *et al.*, 2008). In general, the TP reaction links the 4th amino acid of one peptide stem, known as the donor strand, to that of the 3rd amino acid of an adjacent peptide stem, or acceptor strand, to form a 3-4 crosslink (Vollmer *et al.*, 2008a). The mechanism of the transpeptidation reaction was initially investigated using class C PBPs, which hydrolyse the terminal D-alanine from the stem pentapeptide rather than transferring the acyl group of the reaction intermediate to an amino group (Rasmussen and Strominger, 1978). Penicillin was found to inhibit a *B. subtilis* carboxypeptidase by irreversibly binding to a serine residue present in the active site (Georgopapadakou *et al.*, 1977), suggesting a role for this residue in catalysis. The role of the active site serine was substantiated by further studies utilising a radiolabelled depsipeptide substrate mimic, which permitted the identification of an acyl-enzyme intermediate (Rasmussen and Strominger, 1978). Subsequent crystallographic studies have confirmed the presence of acyl-enzyme complexes formed between PBPs and β -lactams (non-exhaustive list of PDBids: 3UPO, 3OCL, 3OCN, 3UPN, 3ZG8). Taken together, these experiments allowed for the following reaction mechanism of transpeptidation to be proposed: (i) the peptidoglycan stem pentapeptide (Ala-D- γ -Glu/Gln-Lys/meso-DAP-D-Ala-D-Ala) is bound by the PBP; (ii) the active site serine of the PBP attacks the carbonyl carbon of the penultimate D-Ala residue, forming an acyl-enzyme intermediate and releasing the terminal D-Ala; (iii) the carboxyl carbon of the acyl-intermediate is then linked to the lysyl N $_{\epsilon}$ of the adjacent peptide stem to form an isopeptide bond (Scheffers and Pinho, 2005; Sauvage *et al.*, 2008) (**Figure 1.8**, overleaf). In some species of bacteria, a poly-amino acid bridge, usually consisting of glycine, alanine, serine and/or aspartate/asparagine, is used to connect spatially distant peptide stems and allow for a highly crosslinked cell wall (Veiga *et al.*, 2006; Vollmer *et al.*, 2008a). When a poly-amino acid bridge is present, the 3rd step of the TP reaction links the carboxyl carbon of the D-Ala of the donor strand to the amino group of the last amino acid of the bridge, with the 1st amino acid of the bridge forming an isopeptide bond with the N $_{\epsilon}$ of the acceptor strand (Schneider *et al.*, 2004; Vollmer *et al.*, 2008a) (**Figure 1.8**, overleaf). The hydrolysis of the terminal D-Ala residue of the pentapeptide provides the energy required to form

the isopeptide bond, and hence shorter stem peptides, resulting from the action of DD-carboxypeptidases, are unable to act as donor strands for 3-4 crosslink formation (Scheffers and Pinho, 2005). The cleavage of the pentapeptide can therefore be used to regulate the amount of crosslinking present within the cell wall.

Not all peptidoglycan that is incorporated into the cell wall becomes crosslinked. Depending on the method used, HPLC or size exclusion, separation of cell wall material of stationary-phase *B. subtilis* digested with cellosyl, a lysozyme that cleaves the sugar backbone of peptidoglycan whilst leaving crosslinks intact, revealed that between 30% and 40% of stem peptides are crosslinked (Atrih *et al.*, 1999). The authors commented that the value obtained from HPLC analysis (33%) is likely the more accurate of the two due to difficulties in quantifying mucopeptides separated by size exclusion. A similar study of *S. pneumoniae*, again using HPLC to separate digested cell wall components, suggested that 40% of stem peptides participate in crosslinking (Garcia-Bustos *et al.*, 1987). Both these values contrast drastically with the amount of crosslinking found in *S. aureus*, with various studies using HPLC to separate digested cell wall material suggesting that between 70-90% of stem peptides are crosslinked (Snowden and Perkins, 1990). The high percentage of crosslinked stem peptides found in *S. aureus* is suggested to result from a lack of carboxypeptidase activity in this organism, with the terminal D-Ala-D-Ala moiety being retained on all non-crosslinked stem peptides, allowing them to act as donor strands (Scheffers and Pinho, 2005). In *B. subtilis*, *S. pneumoniae* and other species, stem peptides that are not crosslinked during incorporation into the cell wall often have the terminal D-alanine residues removed by the action of DD- and LD-carboxypeptidases, with the major crosslinked species found in digested and HPLC separated *B. subtilis* 168 cell wall being tetrapeptide-tripeptide (Scheffers and Pinho, 2005; Vollmer *et al.*, 2008b; Hoyland *et al.*, 2014).

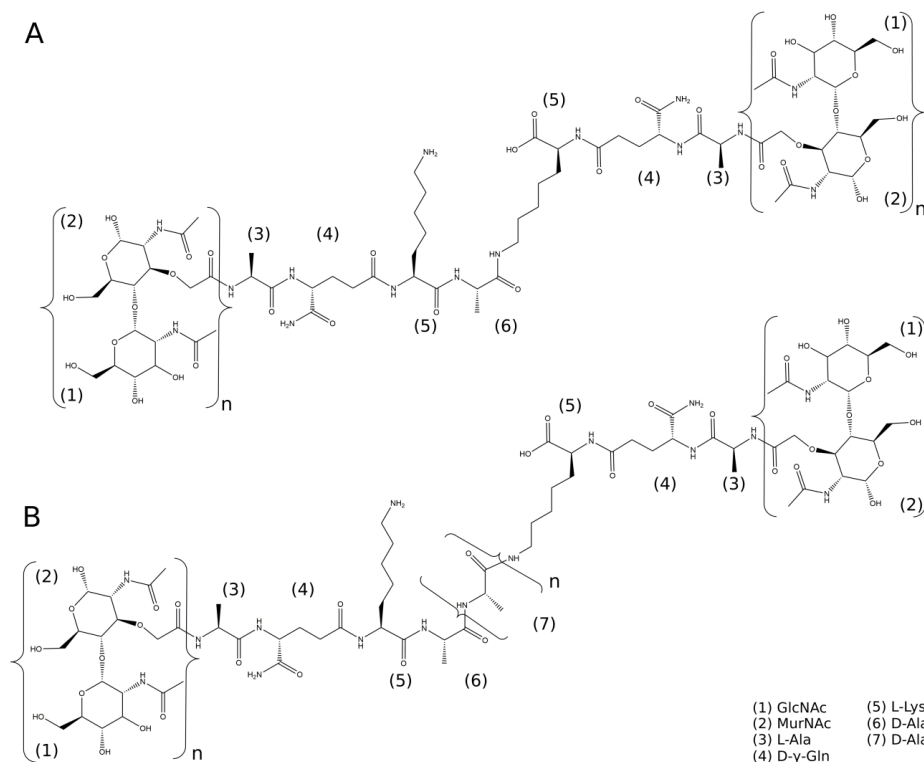


Figure 1.8: Chemical schematic of the cell wall crosslinks. (A) A direct 3-4 crosslink between a stem tripeptide and a stem tetrapeptide, both containing lysine. (B) A crosslink mediated by a peptide crossbridge consisting of D-alanine. This figure was produced using ChemBioDraw v14.0.0.117 (PerkinElmer).

1.5 Localisation of the HMW PBPs

1.5.1 Localisation patterns

One function of the bacterial cell wall is to provide shape to the cell. Several cell morphologies have been documented, and it has been noted that some bacteria are pleomorphic and are able to alter their size and shape in response to environmental conditions. For example, *Lactococcus lactis* has been shown to differentiate into two populations when in a clinical biofilm, with cells at the base of the biofilm retaining the archetypical ovoid shape whilst those at the surface of the biofilm become rod-like, presumably to better exploit the different conditions of the biofilm (Perez-Nunez *et al.*, 2011). Additionally, filamentation in some bacterial species, such as *E. coli*, is thought to allow for better adhesion to a surface, as the surface area contact is increased dramatically (Jens *et al.*, 2013). Thus, under normal circumstances, the cell wall is vital in providing a competitive advantage and ensuring the survival of the bacterial cell. As such, the synthesis of the cell wall must be linked to the bacterial cell cycle; the cell cannot grow without new peptidoglycan being synthesised, and new cell wall must be

deposited as cells divide to ensure a continuous cell wall surrounds both daughter cells. Consequently, it would be expected that the PBPs that incorporate new cell wall material would tend to localise to areas of growth and/or division (Scheffers and Pinho, 2005).

In many rod-shaped bacteria, such as *E. coli* and *B. subtilis*, cells grow by the addition of peptidoglycan on the lateral axis, with division occurring at the mid-point of the cell (**Figure 1.9**, overleaf) (Scheffers and Pinho, 2005). PBPs would therefore be expected to localise to both sites to facilitate the incorporation of new peptidoglycan. This is indeed the case; functional GFP-fusions of the HMW PBPs not involved in sporulation in *B. subtilis* revealed distinct localisation patterns during vegetative growth (Scheffers *et al.*, 2004). PBP1, PBP2c, PBP4 and PBP2b were found to localise to the septum, whilst PBP2a, PbpH, PBP3 and YrrR were found to localise either evenly along the membrane or as foci at the periphery of the cell (Scheffers *et al.*, 2004). Interestingly, a fluorescence signal corresponding to GFP-labelled PBP1 was also observed along the cylinder of the cell, suggesting a dual role for this PBP (Scheffers *et al.*, 2004). Indeed, PBP1 has since been shown to shuttle between the cylinder of the cell and the division site and is believed to be involved in the synthesis of peptidoglycan at both sites (Claessen *et al.*, 2008). In *E. coli*, it has been shown that the deletion of PBP2 results in the appearance of spherical cells, linking the function of PBP2 to elongation. By contrast, mutation of PBP3 inhibited successful division at certain temperatures and produced filaments, indicating that PBP3 is required for cell division (Spratt, 1975). PBP3 was later shown to localise specifically to the septum (Mercer and Weiss, 2002), further supporting its classification as a cell division PBP. Interestingly, in addition to localising in spot-like patterns along the entire bacterial cell, GFP-PBP2 was also found to localise to the septum, suggesting a possible bifunctional elongation and division role for this PBP (Den Blaauwen *et al.*, 2003). Inhibiting PBP2 has been shown to eventually disrupt division (Vinella *et al.*, 1993) because of the swelling of the cell, increasing its diameter and preventing the formation of the Z-ring, the multiprotein division complex that is required for division, rather than as a direct result of the absence of PBP2 from the septum (Vinella *et al.*, 1993).

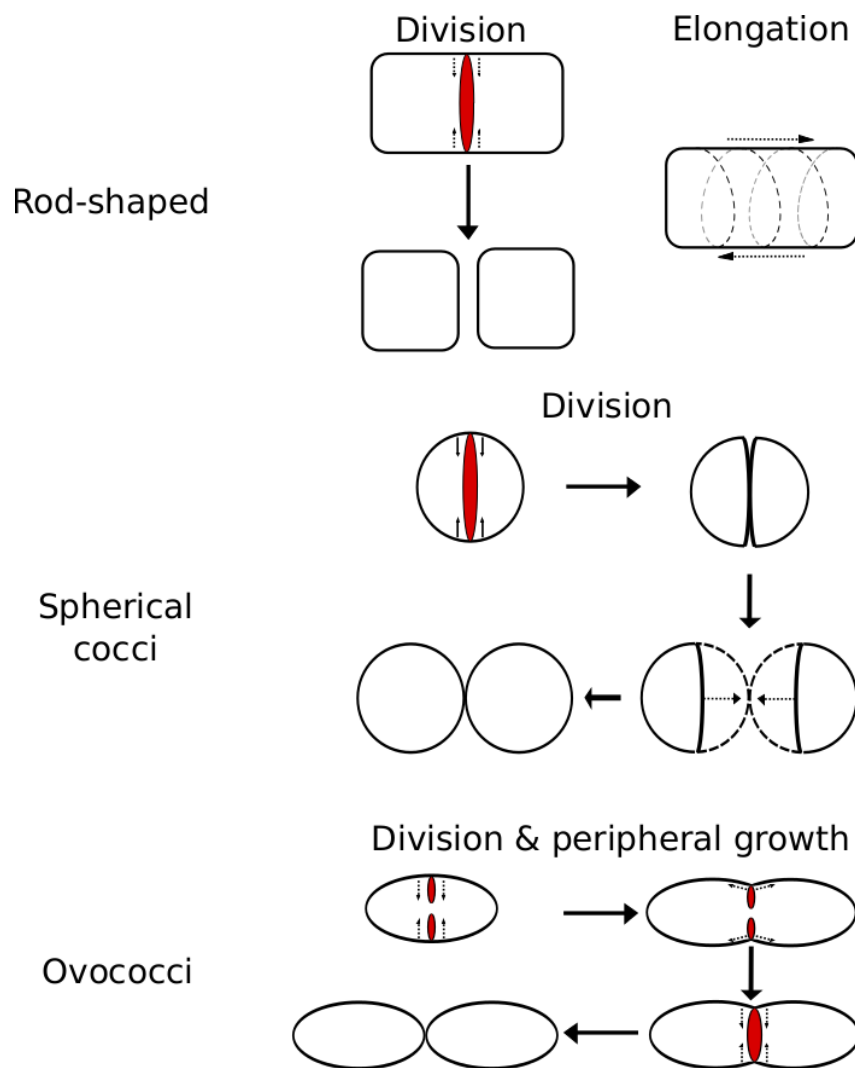


Figure 1.9: Schematic representation of the growth and division methods employed by rod-shape, spherical cocci and ovococci cells. In rod-shaped cells, elongation and division, although linked, occur independently. Spherical cocci grow through division only, with daughter cells 'inflating' after division into spheres. Ovococci experience a form of elongation known as peripheral growth, whereby they elongate during cell division, before the septum has fully closed. The septum is indicated by a red oval, with dotted arrows showing the movement of the cell wall.

Unlike rod-shaped bacteria, spherical cocci such as *Staphylococcus aureus* only incorporate peptidoglycan at the division septum (**Figure 1.9**) (Scheffers and Pinho, 2005). PBPs of *Staphylococcus aureus* would therefore be expected to localise only to the division septum, with the majority of the cell wall consisting of inert peptidoglycan from previous cell divisions. Indeed, PBP2, the only PBP of *S. aureus* that possesses both GT and TP activities (Scheffers and Pinho, 2005), has been shown, through the use of an N-terminally GFP tagged fusion integrated into the chromosome at the native PBP2 locus, to localise to the septum throughout division, being found along the entire septal disk following the closure of the septum (Pinho and Errington, 2005).

In addition to labelling the PBPs with fluorescent reporters (i.e. GFP), sites of peptidoglycan incorporation can be elucidated by labelling the nascent peptidoglycan with fluorescent molecules that interact with the D-Ala-D-Ala moiety of stem pentapeptides, such as fluorescently-labelled vancomycin, as the terminal D-Ala is lost from mature cell wall material as a result of the transpeptidation reaction or through the action of carboxypeptidases (Vollmer *et al.*, 2008a; Vollmer *et al.*, 2008b). Unfortunately, the lack of carboxypeptidase activity in the cell wall of *S. aureus* results in the D-Ala-D-Ala moiety remaining intact in old cell wall material, causing fluorescent vancomycin to stain uniformly (Lunde *et al.*, 2010). Despite this, Lunde *et al.*, (2010) labelled the site of nascent peptidoglycan incorporation in *S. aureus* by using telavancin, a fluorescent vancomycin derivative that binds preferentially to the D-Ala-D-Ala moiety of the lipid II precursor before it undergoes incorporation into the wall. Telavancin was found to stain exclusively at the septum, indicating that lipid II, and therefore peptidoglycan, incorporation occurs only at this site in *S. aureus* (Lunde *et al.*, 2010).

Besides spherical cocci, ovoid cocci (ovococci) are also found in Nature, as exemplified by *Streptococcus pneumoniae*. Whereas spherical cocci synthesise peptidoglycan only at the division septum, ovococci undergo a form of elongation known as peripheral growth (**Figure 1.9**, previous page) (Higgins and Shockman, 1970; Beilharz *et al.*, 2012). These two distinct peptidoglycan insertion events are probably responsible for the characteristic ovoid shape from which the ovococci draw their name (Beilharz *et al.*, 2012). This form of growth was first confirmed by Higgins and Shockman (1970) in *Streptococcus faecalis*, although previous immunofluorescence work by Cole and Hahn (1962) had eluded to it. Higgins and Shockman (1970) fixed a mid-exponential culture of *S. faecalis* and examined the cells by electron microscopy following staining with uranyl acetate and lead citrate. They simulated a time-lapse sequence and determined that cell division was initiated by the formation of the septum at the equatorial ring, with septal cell wall synthesis occurring at the same time. However, Higgins and Shockman (1970) noted that the closure of the septum was halted prematurely and peptidoglycan was inserted at the periphery of the septum, elongating the new hemispheres of the soon-to-be daughter cells. Once elongation had produced new hemispheres of equal size to the old hemispheres, septal cell wall synthesis appeared to be resumed and division proceeded to completion (Higgins and Shockman, 1970;

Beilharz *et al.*, 2012). As with rod-shaped bacteria, the presence of two distinct peptidoglycan incorporation events suggest that PBPs of *S. pneumoniae* localise as two groups; one localises to the septum and forms the septal cell wall, and another localises peripherally to the septum and synthesises the new hemispheres. Immunofluorescence studies using antibodies raised against the extracellular domains of the 5 HMW PBPs found in *S. pneumoniae* (PBP2x, PBP1a, PBP1b, PBP2a and PBP2b) revealed three distinct patterns of localisations at two sites (Morlot *et al.*, 2003). PBP2x and PBP1a were found to localise at the septum whilst PBP2a and PBP2b localised to the peripheral zone of growth. PBP1b was found at both sites, although never at the two simultaneously (Morlot *et al.*, 2003). Thus, in *S. pneumoniae*, two PBPs are utilised for elongation and two for division, with PBP1b capable of participating in either elongation or division.

It is interesting to note that bacteria often also change shape during stationary growth, becoming smaller and rounder, likely resulting from a change in peptidoglycan metabolism (Nystrom, 2004). Recent studies have shown that this appears to directly relate to the amount of D-amino acids produced by the bacteria, with the deletion of *bsrV*, the gene encoding a non-specific racemase capable of catalysing the conversion of a range of L-amino acids into their D-isomers (Espaillat *et al.*, 2014), causing a doubling in thickness of the cell wall (Lam *et al.*, 2009). The D-amino acids, notably D-Met, D-Leu, D-Val, D-Ile, D-Phe and D-Tyr, have been shown by HPLC analysis of digested cell wall material to be incorporated into the cell wall and appear to provide stationary phase bacteria with resistance against osmolysis, suggesting that their incorporation increases the strength of the bacterial cell wall (Lam *et al.*, 2009).

In addition to being incorporated into the cell wall, the production of D-amino acids and their release into the environment appears to be utilised as a form of communication in some species; the addition of D-alanine to media containing *Bacillus* spores prevents their germination. This activity of D-alanine is in direct conflict with the addition of L-alanine, which stimulates germination (Cava *et al.*, 2011). It is thought that, whilst the presence of L-alanine in growth media suggests an abundance of nutrients and thus is supportive of germination, the presence of D-alanine suggests conditions that have caused the lysis and breakdown of the bacterial cell, releasing D-alanine from the cell wall (Cava *et al.*, 2011). Furthermore, D-amino acids have been

shown to be capable of causing the breakup of *B. subtilis* biofilms, causing the detachment of amyloid fibers that hold the biofilm together (Kolodkin-Gal *et al.*, 2010). Presumably the D-amino acids are again acting as messengers, indicating to the bacterial population that cell lysis is occurring and the environment is no longer able to support the biofilm.

1.5.2 Determinants of localisation

Historically, the HMW PBPs were thought to localise through protein-protein interactions, either with other PBPs or with cytoskeletal proteins that function to coordinate the cell cycle (Scheffers and Pinho, 2005). For example, PBP3 of *E. coli*, a monofunctional TPase, has been shown to localise to the septum during division through interactions with FtsW, a lipid II flippase, since PBP3 does not localise in FtsW-depleted cells (Mercer and Weiss, 2002). Di Lallo *et al.*, (2003) used a modified form of the bacterial two-hybrid screen, adapted to prevent non-specific hydrophobic interactions between membrane proteins, to demonstrate that FtsW interacted directly with PBP3 (Di Lallo *et al.*, 2003), whilst Pastoret *et al.*, (2004) and Derouaux *et al.*, (2008) determined that the interaction occurred via the periplasmic loop located between transmembrane helices 9 and 10 (loop 9/10) on FtsW and the first 41 amino acids of PBP3 (Pastoret *et al.*, 2004; Derouaux *et al.*, 2008). In addition, monoclonal antibodies raised against the C-terminus of *E. coli* PBP1B have shown that the enzyme localises to the septum during cell division, with its localisation dependent on the presence of PBP3 (Bertsche *et al.*, 2006). Bacterial two-hybrid studies have suggested a direct interaction between PBP1b and FtsW (Fraipont *et al.*, 2011), however, PBP1b was not detected in the FtsW-PBP3 complex following a co-immunoprecipitation assay, although stable FtsW-PBP3 complexes were detected (Fraipont *et al.*, 2011). While it is possible that the interaction between FtsW and PBP1b is transient or weak, it must also be considered that the interaction observed in the bacterial two-hybrid assay is due to the FtsW-PBP3 and PBP3-PBP1b interactions bringing PBP1b and FtsW into close enough proximity to allow the reporter, the two domains of *Bordetella pertussis* adenylate cyclase (Karimova *et al.*, 1998), to interact and activate transcription of β -galactosidase.

Despite the clear interactions between PBP3 and FtsW, PBP3 does not localise to the division septum when the cell is exposed to furazlocillin, a specific β -lactam inhibitor of

PBP3 that blocks transpeptidation and prevents it from interacting with peptidoglycan stem peptides, suggesting that PBP3's ability to bind substrate is more important than its interaction with FtsW in determining correct localisation (Wang *et al.*, 1998). Further evidence of substrate-dependent localisation of the PBPs can be found in a range of bacteria. In *B. subtilis*, functional GFP-fusions of PBP2a and PbpH, class B PBPs required for cell elongation (Wei *et al.*, 2003), have recently been shown to be sequestered into patches along the lateral membrane when cells are exposed to nicin, a lantibiotic that sequesters lipid II away from its physiological location through interactions with the pyrophosphate moiety, indicating that the availability of nascent peptidoglycan determines their localisation (Lages *et al.*, 2013).

Substrate recognition has also been shown to affect HMW PBP localisation in *S. aureus*; here, the β -lactam antibiotic oxacillin caused the delocalisation of PBP2 from the division septum, but did not affect other Z-ring proteins or lipid II localisation (Pinho and Errington, 2005). However, oxacillin-exposed PBP2 did localise to the septum in the presence of PBP2A, a β -lactam resistant PBP, suggesting a combination of protein-protein interactions and substrate recognition localise PBP2 in *S. aureus* (Pinho and Errington, 2005).

1.6 Transpeptidation substrate availability is controlled by carboxypeptidases

Carboxypeptidases are hydrolytic enzymes that cleave the C-terminal residue from peptides. In the context of the cell wall, this involves the cleavage of the stem peptide that is found on the MurNAc moiety (Vollmer *et al.*, 2008b). The carboxypeptidases of the cell wall are a very diverse group of enzymes that contains members capable of cleaving progressively along the stem peptide; DD-carboxypeptidases cleave the terminal D-Ala from the stem pentapeptide, LD-carboxypeptidases remove the now-terminal D-Ala from the tetrapeptide and DL-carboxypeptidases and LD-carboxypeptidases subsequently catalyse the formation of the dipeptide and mono peptide respectively (Vollmer *et al.*, 2008b) (**Figure 1.10**). The final enzymes, which cleave the L-alanine from the MurNAc sugar, are known as muramidases rather than carboxypeptidases (Vollmer *et al.*, 2008b) and will not be considered further.

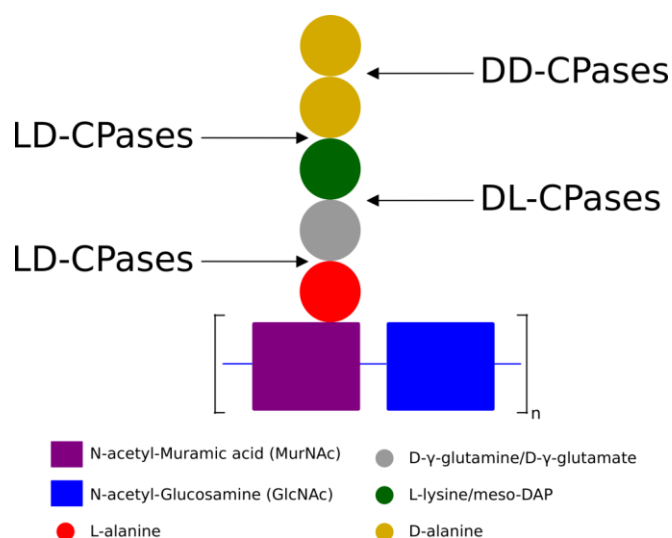


Figure 1.10: Schematic showing the cleavage points of various cell wall carboxypeptidases. CPase = carboxypeptidase; DD = between two D-amino acids, LD = between the C-terminus of an L-amino acid, and the N-terminus of a D-amino acid; DL = between the C-terminus of a D-amino acid, and the N-terminus of an L-amino acid.

The ability of carboxypeptidases to remove the terminal D-Ala-D-Ala moiety from peptidoglycan allows them to control substrate availability for transpeptidation, implicating them in controlling the localisation of the HMW PBPs and thus the morphology of the bacterial cell. Direct evidence of DD-carboxypeptidases controlling the availability of the transpeptidation substrate can be found in DacA, a low molecular weight (LMW) PBP from *S. pneumoniae* (**Figure 1.11**).

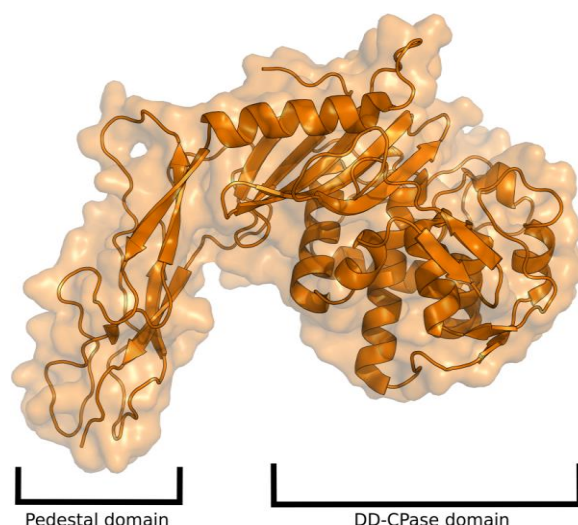


Figure 1.11: The crystal structure of a DacA (PBP3) from *S. pneumoniae*, a class C PBP. DacA contains a DD-carboxypeptidase (DD-CPase) domain, with a second domain similar to the pedestal domain of class B PBPs, likely positioning the PBP to be within close proximity of its substrate (Morlot *et al.*, 2005). DacA is shown as an orange cartoon ribbon with transparent surface.

In wild-type cells, immunofluorescence studies demonstrated that DacA was found to localise over the entire surface of *S. pneumoniae*, with the exception of the equatorial division site (Morlot *et al.*, 2004). Conversely, immunofluorescence was used to show that the HMW PBPs, PBP2x and PBP1a localised only to the Z-ring, the multicomponent division complex at the site of division, consistent with their role as division PBPs (Morlot *et al.*, 2003). In agreement with the observed localisation patterns of DacA, PBP2x and PBP1a, staining of *S. pneumoniae* with fluorescent vancomycin demonstrated that the stem pentapeptide was located primarily at the division site (Morlot *et al.*, 2004). Upon loss of DacA activity, the ordered localisation of PBP2x and PBP1a to the Z-ring was abolished, whilst other proteins that form the Z-ring were unaffected (Morlot *et al.*, 2004). Severe morphological defects were also observed, with cells undergoing asymmetric division and becoming smaller and more spherical (**Figure 1.12**) (Schuster *et al.*, 1990; Morlot *et al.*, 2004; Barendt *et al.*, 2011). Barendt *et al.*, (2011) were unable to repeat the localisation pattern observed for DacA by Morlot *et al.*, (2004), however, Barendt *et al.*, (2011) used FLAG-tagged DacA in their immunofluorescence studies, rather than raising antibodies against DacA as did Morlot *et al.*, (2004), raising the possibility of aberrant localisation of the FLAG-tagged protein or non-specific binding of the anti-FLAG antibody. Whilst these possibilities remain, the clear differences between the localisations observed by Morlot *et al.*, (2004) and Barendt *et al.*, (2011) will have to be reconciled if we are to fully understand the role of DacA in cell morphogenesis.

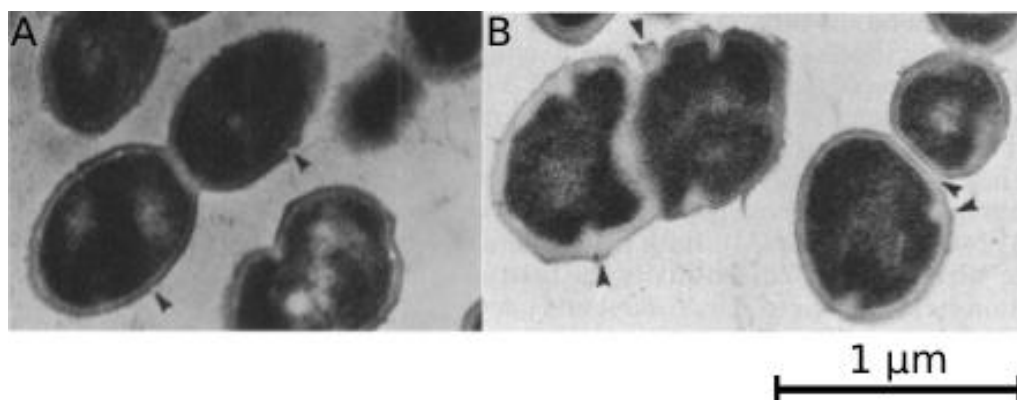


Figure 1.12: Electron microscopy showing the morphologies of wild-type *S. pneumoniae* (A) and a $\Delta dacA$ strain (B), taken from Schuster *et al.*, (1990). $\Delta dacA$ *S. pneumoniae* divide asymmetrically and produce mis-shapen, rounder cells of varying sizes. Arrows indicate division septa.

Several carboxypeptidases have been shown to be essential for maintaining cell morphology in the helical bacteria *Helicobacter pylori* and *Campylobacter jejuni*. Although HMW PBPs have been identified in both species, LMW PBPs that provide carboxypeptidase functionalities are surprisingly absent, and have been replaced by non-PBP carboxypeptidases. The first carboxypeptidase discovered to affect the morphology of *H. pylori* was Csd3/HdpA, which possesses both DD-carboxypeptidase and DD-endopeptidase activities and is a zinc metallopeptidase (Bonis *et al.*, 2010). The physiological role of HdpA was elucidated by deletion studies, which demonstrated that $\Delta hdpA$ strains of *H. pylori* formed shorter, wider cells than the wild-type, with a proportion of the population forming a branched phenotype (**Figure 1.13**) (Bonis *et al.*, 2010). When *hdpA* was placed under the control of an inducible promoter and overexpressed, HdpA resulted in cells becoming spherical, but were still able to divide as normal (Bonis *et al.*, 2010). Peptidoglycan analysis of cells overexpressing HdpA revealed a decrease in the peptidoglycan pentapeptide content of the cell wall, confirmed by fluorescent vancomycin labelling of the pentapeptide in growing cells (Bonis *et al.*, 2010).

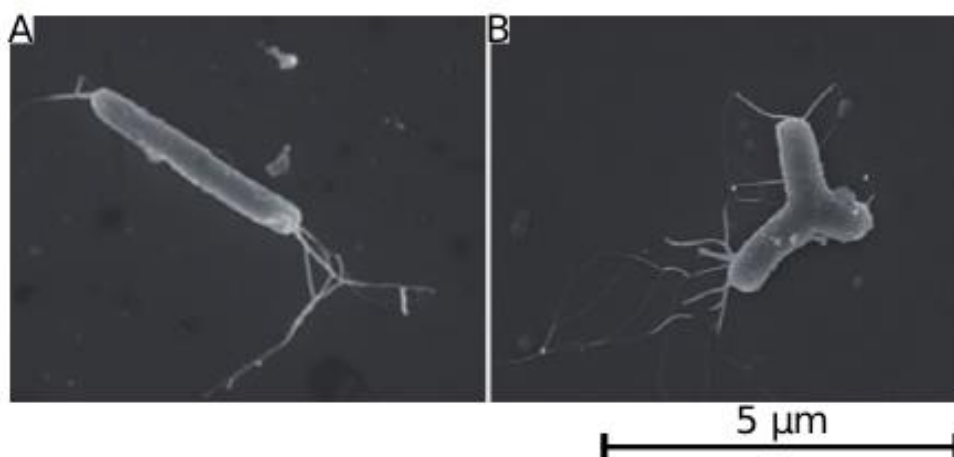


Figure 1.13: Scanning electron microscopy pictures of N6 *H. pylori* (A) and $\Delta hdpA$ N6 *H. pylori* (B), taken from Bonis *et al.*, (2010). The HdpA mutant shows a branched phenotype.

Pgp2 (*C. jejuni*) and Csd6 (*H. pylori*) also contribute to cell shape maintenance; deletion of either caused helical-to-rod transitions (Sycuro *et al.*, 2013; Frirdich *et al.*, 2014). Sequence similarities suggested that these enzymes may be LD-transpeptidases, but Sycuro *et al.*, (2013) noted that Costa *et al.* (1999) failed to identify LD-crosslinks in the

peptidoglycan of *H. pylori* and suggested that Csd6 is, instead, an LD-carboxypeptidase (Sycuro *et al.*, 2013). The function of Csd6 was confirmed when the protein was purified from heterologous overexpression in *E. coli* and incubated with the peptidoglycan of a $\Delta csd1csd6$ mutant of *H. pylori*. It was discovered that Csd6 was able to completely hydrolyse the uncrosslinked tetrapeptides present in the peptidoglycan, but was unable to cleave the terminal D-alanine from crosslinked stem tetrapeptides (Sycuro *et al.*, 2013). Unlike the loss of the pentapeptide through DD-carboxypeptidase activity, which prevents crosslink formation as only the pentapeptide can act as a donor in the transpeptidation reaction, it was not immediately apparent why the loss of the tetrapeptide would so drastically affect the morphology of the cell as both the stem tetrapeptide and the tripeptide could still act as acceptors for the TP reaction. Thus a decrease in the amount of either the tetrapeptide or tripeptide content of the cell, together with an increase in the amount of the other should not affect the cell in any meaningful way (Sycuro *et al.*, 2013). A similar study of Pgp2 revealed the same results (Firdich *et al.*, 2014), allowing both proteins to be classified as LD-carboxypeptidases.

1.7 Other carboxypeptidase functions

In addition to controlling the availability of the D-Ala-D-Ala moiety, and thereby controlling the localisation of the HMW PBPs, carboxypeptidases perform a number of other functions within the bacterial cell, being required for the recycling of peptidoglycan, for instance, and in the resistance against antibacterial agents.

1.7.1 Peptidoglycan recycling

Although peptidoglycan is often considered a static structure, the cell wall is constantly being degraded and re-synthesised, with a turnover rate of around 50% per generation in many species (Reith and Mayer, 2011). Radiolabelled meso-DAP has been used to demonstrate that only 6-8% of the cell wall material of *E. coli* is lost to the environment, despite the turnover of half the cell wall per generation (Goodell and Schwarz, 1985). The much greater mass of peptidoglycan in the cell wall of Gram-positive bacteria in comparison to Gram-negatives implies that is the former group must recycle the soluble products of peptidoglycan degradation or else face a significant loss of resource. Recently, a number of the genes required for recycling in

E. coli have been identified in Gram-positive organisms, however, several key enzymes required for turnover in *E. coli* have not been found in Gram-positives, including AmpG, required for the uptake of anhydromuropeptides (Reith and Mayer, 2011). Whilst it is not clear whether cell wall recycling systems or the chemical compositions of mature cell walls is the root of the genetic differences between the two bacterial classes, fractionation of the cell wall of *B. subtilis* by HPLC has revealed that anhydroMurNAc only makes up 0.4% of all muropeptides in this species and thus are likely not targeted for recycling, possibly explaining the absence of AmpG (Atrih *et al.*, 1999; Reith and Mayer, 2011).

For peptidoglycan recycling to occur, the mature cell wall must first be degraded. This occurs by the action of a number of enzymes, including endopeptidases that cleave the peptide crosslinks and lytic transglycosylases that hydrolyse the β 1-4 links of the glycan strands (Reith and Mayer, 2011; Lee *et al.*, 2013). The predominant form of muropeptide found in the cell wall of *E. coli* following digestion with a muramidase from *Chalaropsis* *sp.* and separation via HPLC was determined to be GlcNAc-MurNAc-tetrapeptide or crosslinked tetrapeptide-tetrapeptide (Glauner *et al.*, 1988). Once inside the cytoplasm, the GlcNAc-MurNAc-tetrapeptide is degraded further by the actions of amidases to produce GlcNAc, MurNAc and the tetrapeptide (Jacobs *et al.*, 1995; Votsch and Templin, 2000; Reith and Mayer, 2011). The importance of hydrolysing the tetrapeptide to produce the tripeptide in the next step of the pathway has been highlighted by work by Templin *et al.* (1999). Deletion studies of LdcA, an LD-carboxypeptidase and member of the S66, clan SS family of carboxypeptidases (Templin *et al.*, 1999; Korza and Bochtler, 2005; Das *et al.*, 2013), resulted in cells that lysed soon after entering the stationary phase. Digestion of the sacculi of these cells by a muramidase and analysis by HPLC revealed that the cells lost the majority of their stem pentapeptides and 3-4 crosslinks, dramatically reducing the amount of crosslinks present within the cell (Templin *et al.*, 1999). This reduction of crosslinking was suggested by Templin *et al.*, (1999) to be due to the inability of the cell to add a single D-alanine to the tetrapeptide to produce the pentapeptide, which is required to act as a donor in the formation of crosslinks. Although LdcA is found in *B. subtilis* and other Gram-positives (Reith and Mayer, 2011), its importance has yet to be demonstrated, as the majority of muropeptides found in the cell wall of *B. subtilis*

were determined by HPLC to be tripeptides (Atrih *et al.*, 1999), and thus can be recycled and used directly in peptidoglycan biosynthesis.

Of all the LD-carboxypeptidases studied to-date, LdcA is perhaps the best characterised. LdcA has been shown to be specific for the meso-DAP-D-Ala (or L-Lys-D-Ala) peptide bond, but only in soluble peptidoglycan fragments; the enzyme was apparently inactive against mature peptidoglycan (Templin *et al.*, 1999; Baum *et al.*, 2005), consistent with its role in peptidoglycan recycling and localisation to the cytoplasm. Owing to the lysis of $\Delta ldcA$ cells upon entering the stationary phase (Templin *et al.*, 1999), LdcA has been considered a potential antibiotic target, with several inhibitors having been identified, including the β -lactam antibiotics, but only at high concentrations (Baum *et al.*, 2005). Presumably, LdcA is inhibited only by high concentrations of β -lactams because it has evolved to recognise the meso-DAP/L-Lys-D-Ala moiety of the stem peptide, rather than the D-Ala-D-Ala moiety that β -lactams mimic. With this in mind, Baum *et al.*, (2005) were successful in developing a potent inhibitor of LdcA, DTZ, that mimics meso-DAP/L-Lys-D-Ala (**Figure 1.14**).

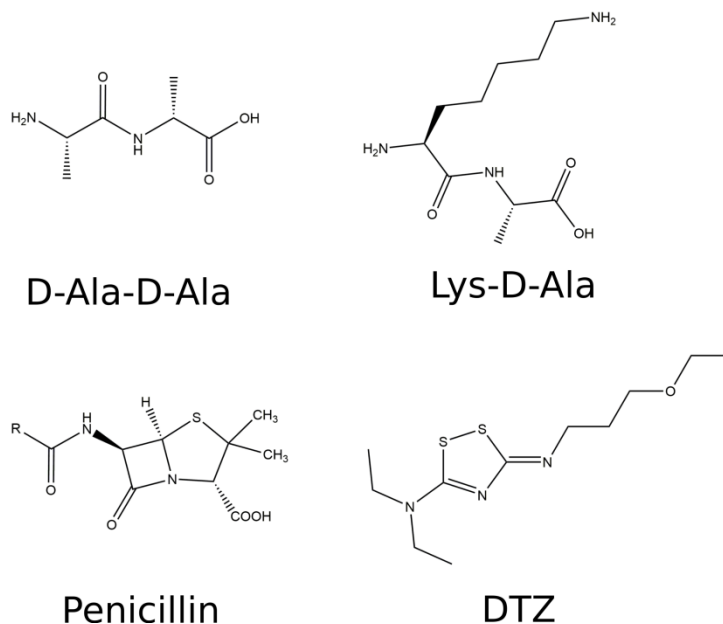


Figure 1.14: Chemical structures of D-Ala-D-Ala (top left) and Lys-D-Ala (top right) and their non-hydrolysable mimics (bottom left and right). Penicillin mimics D-Ala-D-Ala but is non-hydrolysable, irreversibly binding to the TPase domain of PBPs. DTZ, described by Baum *et al.*, (2005), mimics Lys-D-Ala and inhibits LdcA. The R group on penicillin represents a variable group. This figure was produced using ChemBioDraw v14.0.0.117 (PerkinElmer).

The crystal structures of LdcA from both *Pseudomonas aeruginosa* (PaLdcA) (PDBid: 1ZL0, 1ZRS, 2AUM and 2AUN) (Korza and Bochtler, 2005) and *Novosphingobium aromaticivorans* (NaLdcA) (PDBid: 3G23) (Das *et al.*, 2013) have been solved; the proteins superpose with a core RMSD of 2.0 Å over 242 Cαs. LdcA is a globular protein comprising two domains (**Figure 1.15**, overleaf). The N-terminal sheet domain comprises a 5-stranded parallel β-sheet, with three α-helices connecting the strands. The C-terminal domain contains a 7-stranded β-barrel, that is composed of a 3-stranded parallel β-sheet and a four-stranded antiparallel β-sheet (Korza and Bochtler, 2005). The active site of LdcA contains a catalytic triad, consisting of Ser104, His261 and Glu191 in NaLdcA (Ser115, His285 and Glu217 in PaLdcA), and is located between the two domains. The location of the active site is similar to the nucleophilic elbow of αβ-hydrolases, in which the active site serine is located in a sharp turn between a β-strand and an α-helix (Korza and Bochtler, 2005). Though a structure of LdcA with bound substrate is not available, docking experiments performed by Das *et al.*, (2013) on NaLdcA have implicated several residues in substrate recognition and binding. Of the residues identified, Tyr78, Asp105, Asn167, and Glu195 of NaLdcB are highly conserved amongst LdcA orthologues, but their essentiality was not assessed through biochemical methods (Das *et al.*, 2013).

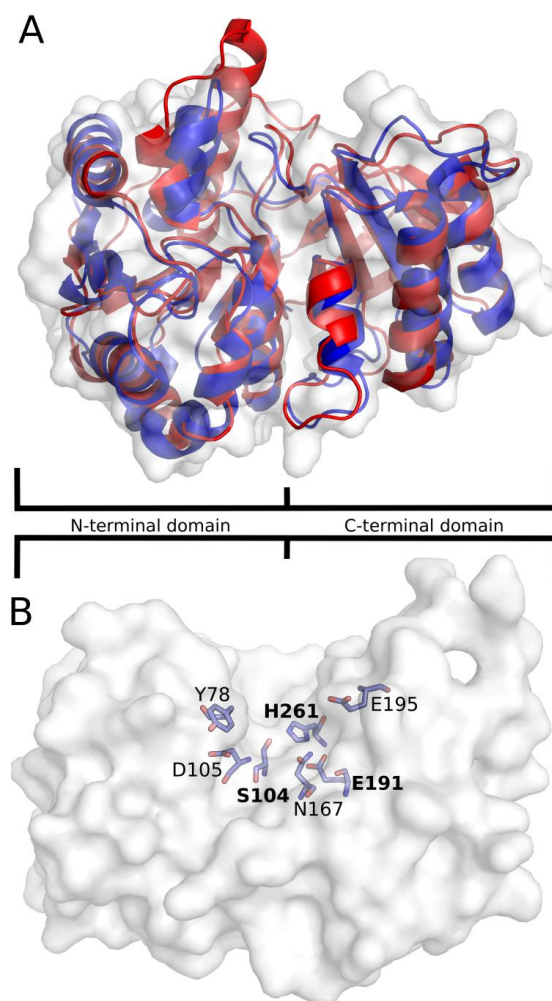


Figure 1.15: (A) The structures of *NaLdcA* and *PaLdcA* superposed; (B) Surface representation of *NaLdcA* with catalytically important residues shown. (A) The two LdcA structures superpose well, with both N- and C- terminal domains present in both structures, although there are some differences in the N-terminal domains. *PaLdcA* is shown as a red cartoon ribbon, *NaLdcA* is shown as a blue cartoon ribbon. The transparent surface is that of *NaLdcA*. (B) The catalytic triad (H261, S104, E191) is located between the N- and C- terminal domains and is shown in stick representation with bold annotation. Residues identified by Das *et al.*, (2013) as important for ligand binding are also shown as sticks but with non-bold annotation. Residue numbering corresponds to *NaLdcB*. The surface of *NaLdcA* is shown in translucent white.

1.7.2 Antibiotic resistance

Whilst the terminal D-Ala-D-Ala moiety of the stem pentapeptide is essential for the formation of the 3-4 peptidoglycan crosslink, which links independent glycan strands together into the peptidoglycan mesh, it is a figurative chink in the armour of the cell wall. Not only do molecules that target D-alanine effectively inhibit the formation of crosslinks and thus kill bacteria, they only affect bacterial cells as D-alanine is not utilised by eukaryotes. A class of antibiotics, the glycopeptides, function in this manner and are exemplified by vancomycin (Yim *et al.*, 2014) (**Figure 1.16**).

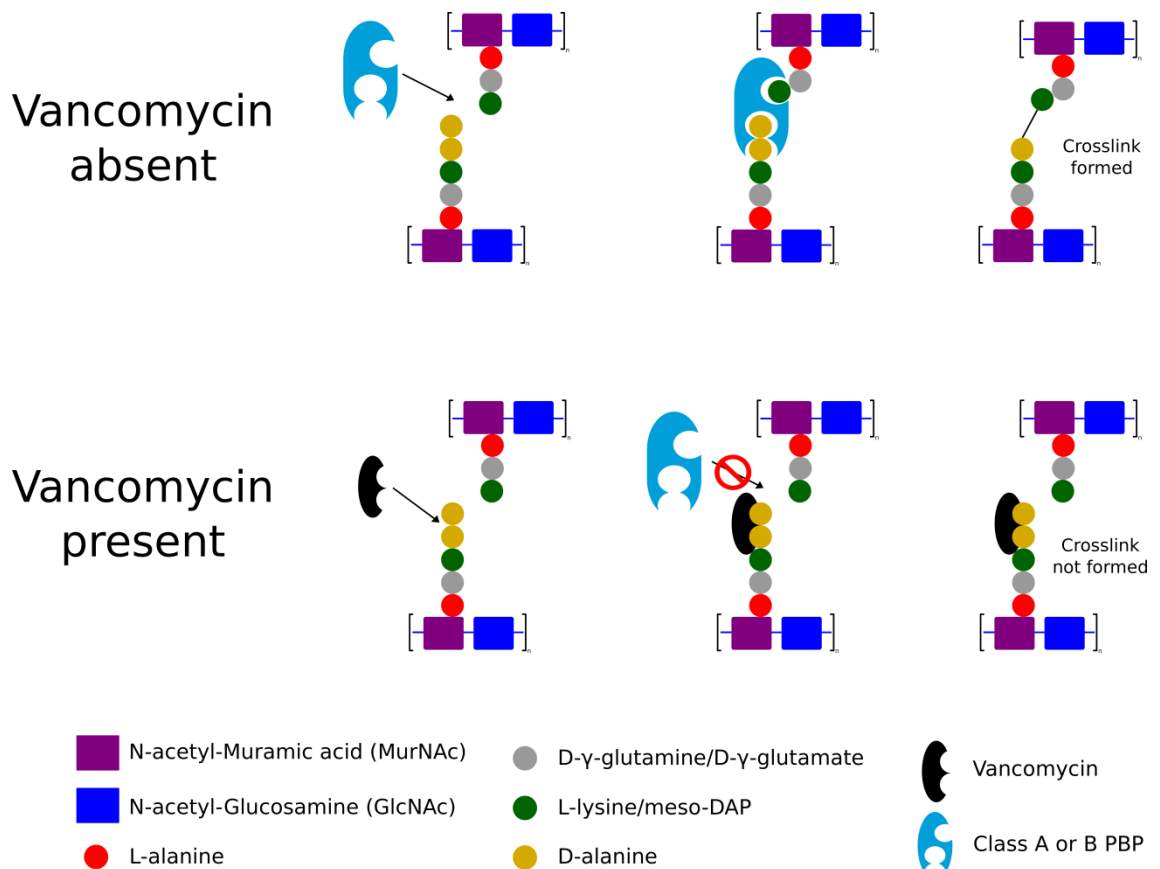


Figure 1.16: Schematic showing the mode of action of vancomycin. Top panel: vancomycin is absent so the PBP is able to bind both the donor and acceptor stem peptides and perform the transpeptidation reaction, forming a crosslink. Bottom panel: vancomycin binds to the terminal D-Ala-D-Ala of the stem pentapeptide, preventing PBPs from binding and inhibiting crosslink formation.

Unfortunately, many bacteria have evolved resistance to vancomycin. Many vancomycin-resistant species have replaced the terminal D-Ala-D-Ala residues of the stem pentapeptide with D-Ala-D-Lac or D-Ala-D-Ser, towards which vancomycin has little or no affinity (Popieniek and Pratt, 1987; Bugg *et al.*, 1991), but which are still recognised as donor stems by the PBPs for use in TP reactions (Arthur *et al.*, 1996). It should be noted, however, that PBP5, required for β -lactam resistance in *Enterococcus faecium* D366, appears to be unable to recognise D-Lac terminating pentapeptides as vancomycin and β -lactams, when used together, are able to lyse this strain (al-Obeid *et al.*, 1992).

In addition to replacing the terminal D-Ala-D-Ala of the stem pentapeptide with an amino acid to which vancomycin cannot bind, vancomycin resistant cells also produce a DD-carboxypeptidase, VanY, that cleaves the terminal D-alanine from stem

pentapeptides. VanY thus reduces the amount of D-Ala-D-Ala present on the cell surface to provide resistance to vancomycin (Wright *et al.*, 1992). Like other vancomycin resistance genes (*vanR*, *vanS*, *vanA*, *vanH* and *vanX*) *vanY* is found on the vancomycin resistance plasmid pIP816 (Wright *et al.*, 1992) and encodes a ~35 kDa protein (Wright *et al.*, 1992) with hydrolytic activity only. VanY does not catalyse the transpeptidation reaction (Wright *et al.*, 1992), is not sensitive to β -lactam antibiotics and does not display any β -lactamase activity, suggesting that its resistance to inhibition by β -lactams does not stem from the hydrolysis of the antibiotic (Wright *et al.*, 1992).

Recently, an enzyme that combines the functions of VanY and VanX, a further protein involved in vancomycin resistance that functions as a D-Ala-D-Ala dipeptidase, has been identified. Known as VanXYc, it is found in organisms that replace the terminal D-Ala-D-Ala of the dipeptide with D-Ala-D-Ser (Reynolds *et al.*, 1999) and is active against both the pentapeptide terminating in D-Ala-D-Ala and the D-Ala-D-Ala dipeptide, but not against stem peptides ending in D-Ala-D-Ser, nor the D-Ala-D-Ser dipeptide (Podmore and Reynolds, 2002). A second enzyme, VanXYg, was identified based upon sequence similarities to VanXYc (McKessar *et al.*, 2000), although the activity of VanXYc against DD-pentapeptides has yet to be demonstrated. The crystal structures of both VanXYg and VanXYc have been solved recently by Meziane-Cherif *et al.*, (2014) and provide molecular insights into VanY-like enzymes.

Both VanXYc and VanXYg are zinc metallopeptidases, confirming previous classifications by the MEROPS database as clan MD, family M15 members (Rawlings *et al.*, 2012). This family contains many enzymes involved in a range of metabolic processes, including other bacterial cell wall modifying enzymes, and has often been referred to as the LAS family, because the family includes *lysostaphin*, a canonical D-Ala-D-Ala dipeptidase and the N-terminal peptidase domain of *sonic hedgehog* (Bochtler *et al.*, 2004). Members of the LAS family are characterised by the presence of a core antiparallel β -sheet formed from at least four β -strands (Bochtler *et al.*, 2004); this core is found in both VanXY enzymes, although two extra β -strands are present, which is not unusual within the family (Bochtler *et al.*, 2004; Meziane-Cherif *et al.*, 2014). Several α -helices (α 1-7) are positioned around the β -sheet and produce a globular, V-shaped domain at the bottom of which is a zinc binding site (**Figure 1.17**).

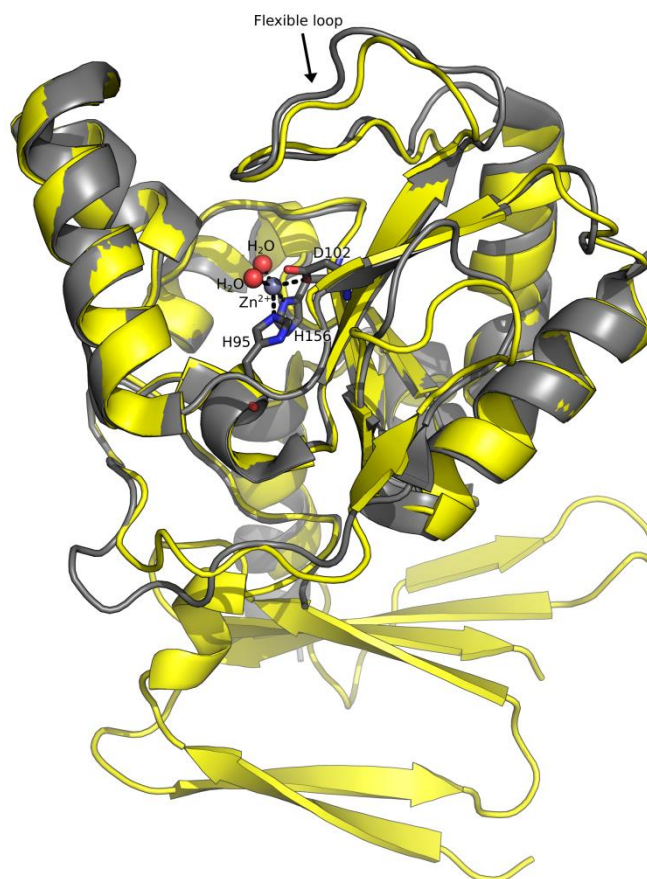


Figure 1.17: The superposed crystal structures of VanXYg and VanXYc. Both structures are shown as cartoon ribbons, with VanXYc shown in grey and VanXYg shown in yellow. The active site of VanXYc is shown, with the catalytic zinc ion shown as a grey sphere, coordinating residues as grey sticks and solvent molecules as red spheres. The active site of VanXYg is identical, with the exception that one the zinc-bound water molecules is replaced by a phosphate. The active site of VanXYg is not shown for clarity. The flexible lid region of both proteins is highlighted.

The zinc is coordinated by two histidines (His95 and His156 in VanXYc; His107 and His167 in VanXYg) and an aspartate (Asp102 in VanXYc and Asp114 in VanXYg), conforming to the zinc-binding consensus sequence found in LAS family members (His- $X_{3/6}$ -Asp- $X_{\text{not conserved}}$ -His) (Bochtler *et al.*, 2004), with two solvent molecules (two waters in VanXYc and a water and a phosphate in VanXYg) acting to complete a trigonal bipyramidal coordination sphere (**Figure 1.17**, previous page). The active site of both enzymes is blocked by a flexible loop region (amino acids 105-121 in VanXYc; 117-132 in VanXYg) that the authors predict changes its conformation to permit the pentapeptide substrate to be accommodated (Meziane-Cherif *et al.*, 2014). In addition to the apo-enzyme, the structures of both VanXY enzymes were solved in complex with a phosphinate transition state mimic of D-Ala-D-Ala, and VanXYc was solved with one of its substrates, D-Ala-D-Ala, bound (**Figure 1.18**) (Meziane-Cherif *et al.*, 2014).

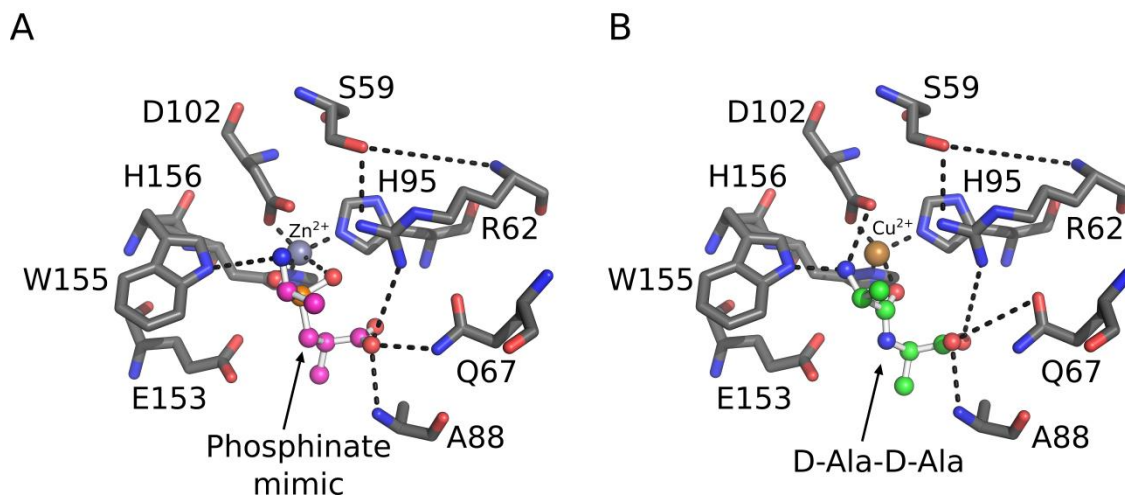


Figure 1.18: The active site of VanXYc with a phosphinate mimic of D-Ala-D-Ala bound (A) and D-Ala-D-Ala bound (B). Both molecules interact with VanXYc in similar ways, with almost identical hydrogen bonds formed. Whilst D-Ala-D-Ala interacts with a copper ion through the carbonyl oxygen of the N-terminal D-alanine whilst the phosphinate mimic interacts through two oxygens, mimicking a tetrahedral transition state. Amino acids are shown as sticks with grey carbons, the zinc as a grey sphere, the copper as an orange sphere and the bound ligands in ball-and-stick representation. D-Ala-D-Ala is shown with green carbons and the phosphinate mimic with pink carbons.

The positions of the D-Ala-D-Ala substrate and the phosphinate mimic were identical and allowed Meziane-Cherif *et al.*, (2014) to identify important interactions, with a range of nearby residues contributing hydrogen bonds (**Figure 1.18**). The C-terminal D-Ala of the dipeptide, and its equivalent in the phosphinate analogue, was found to form hydrogen bonds with Gln67, Ala88 and Arg62 (Gln79, Ala100 and Arg74 in VanXYg); amino acids that excluded substrate residues larger than alanine due to steric clashes. The N-terminal D-Ala of the D-Ala-D-Ala dipeptide was observed to coordinate the metal ion via its carbonyl oxygen, polarising the bond and placing the peptide bond in an ideal position to be hydrolysed by a nucleophilic water activated by the nearby Glu153 (Glu164 in VanXYg). The phosphinate mimic coordinated to the zinc ion with two oxygens and so likely mimicked a tetrahedral reaction intermediate (Meziane-Cherif *et al.*, 2014). Interestingly, Trp155 (Trp166 in VanXYg) appeared to aid in coordinating the dipeptide, but in doing so blocked the active site from the external solvent. Whilst this is the ideal environment for the hydrolysis of a dipeptide, this orientation of Trp155 would exclude the pentapeptide substrate of VanXYc and so presumably moves to accommodate the larger ligand, although evidence for this has yet to be observed.

1.8 LdcB/DacB

An LD-carboxypeptidase, LdcB (previously DacB) was first identified in *L. lactis* (Courtin *et al.*, 2006). The separation of digested cell wall material from an *ldcB*-null mutant by reverse-phase HPLC and subsequent analysis by MALDI-TOF mass spectrometry revealed an increase in the amount of tetrapeptide stem peptides found within the cell wall compared to the wild-type, suggesting that *ldcB* functioned as an LD-carboxypeptidase (Courtin *et al.*, 2006). An increase in the tetrapeptide content of cell walls from *ldcB* deficient strains has also been observed more recently in *B. subtilis* and *S. pneumoniae* (Barendt *et al.*, 2011; Abdullah *et al.*, 2014; Hoyland *et al.*, 2014). The $\Delta ldcB$ mutants of *L. lactis* did not suffer from growth defects under laboratory conditions, however cell morphology was not assessed (Courtin *et al.*, 2006). As with *L. lactis*, growth defects were not observed in *B. subtilis* $\Delta ldcB$ strains under laboratory conditions; by contrast, *ldcB* null mutants of *S. pneumoniae* exhibit severe division asymmetry, producing small, round cells or long, thick cells as a result of random septum localisation (**Figure 1.19**, overleaf) (Barendt *et al.*, 2011).

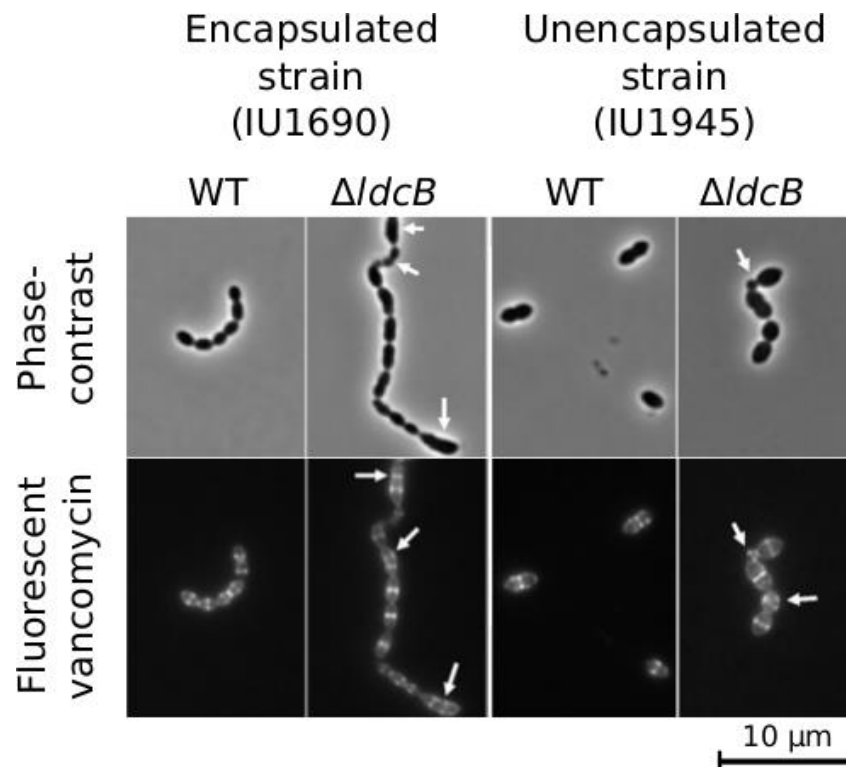


Figure 1.19: Phase-contrast and fluorescent vancomycin images of wild-type and $\Delta ldcB$ encapsulated (IU1690) and unencapsulated (IU1945) *S. pneumoniae*, taken from Barendt *et al.*, (2011). In both strains, the deletion of *ldcB* results in the appearance of small, round cells or long, thick cells (indicated by arrows in the upper panels) and a result of the formation of multiple division septa (indicated by arrows in the lower panels).

Barendt *et al.*, (2011) have suggested that LdcB acts on stem tetrapeptides to localise PBPs to the division septa, in a manner similar to other carboxypeptidases (Section 1.5). This suggestion is consistent with the appearance of multiple division septa in *ldcB*-null mutants (Figure 1.19), but is inconsistent with the localisation of LdcB in patches all over the surface of *S. pneumoniae* (Figure 1.20, overleaf). This inconsistency can perhaps be reconciled as LdcB requires DacA to function (Barendt *et al.*, 2011); although LdcB is present at the septum, it would not act on uncrosslinked stem pentapeptides without DacA, which is absent from the division site, having processed stem pentapeptides (Morlot *et al.*, 2004). Alternatively, the presence of LdcB foci at the division septum in a small population of cells (Figure 1.20, overleaf) may be due to presence of a 3x FLAG epitope tag, which allows the visualisation of LdcB but possibly influences its localisation. It is unclear if LdcB activity influences the localisation of the PBPs, as no localisation studies have been performed in $\Delta ldcB$ cells. Alternatively, LdcB activity could influence the localisation of other cell division proteins that localise through their interactions with peptidoglycan, such as StkP, which is thought to link cell division and peripheral cell growth (Beilharz *et al.*, 2012).

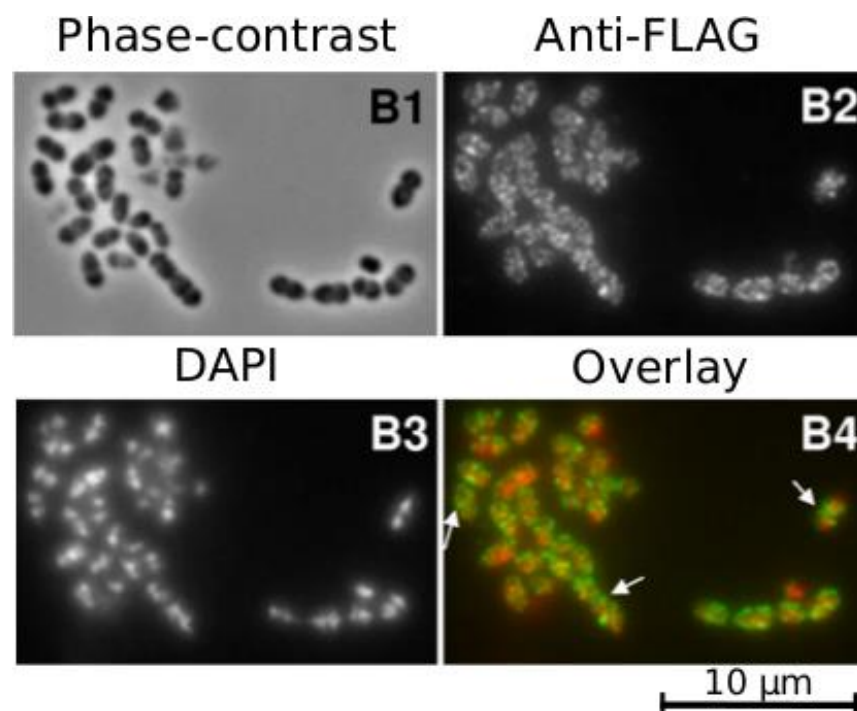


Figure 1.20: The localisation pattern of 3xFLAG tagged LdcB, taken from Barendt *et al.*, (2011). Cells are unencapsulated *S. pneumoniae* growing in the exponential phase. A polyclonal anti-flag antibody was used to visualise LdcB-3xFLAG (B2), DAPI was used to stain nucleoids (B3) and phase-contrast microscopy used to obtain an outline of the cell (B1). A false-colour image (B4) was produced, with DAPI coloured red and LdcB-3xFLAG coloured green. Arrows indicate the presence of LdcB-3xFLAG at the septum in some cells.

Finally, despite several published studies on LdcB, its LD-carboxypeptidase activity has not been demonstrated directly, though the presence of increased amounts of tetrapeptide in $\Delta ldcB$ strains implies this activity. Sequence similarities with VanY and the VanXY enzymes suggest that LdcB adopts a similar, zinc metallopeptidase fold (Barendt *et al.*, 2011), but neither the structure of LdcB, nor of the VanXY enzymes, were known at the outset of this study.

1.9 Aims of this study

In order to confirm the classification of LdcB as an LD-carboxypeptidase, and to elucidate its structure, LdcB from both *B. subtilis* (*BsLdcB*) and *S. pneumoniae* (*SpLdcB*) were studied. This thesis details the overexpression of *SpLdcB* and *BsLdcB* in a heterologous host (*E. coli*) and their subsequent purification to homogeneity. To determine the substrate specificity of LdcB, the activity of purified proteins was assessed against a range of synthetic substrates (**Section 4.3**). To analyse the LD-carboxypeptidase designation of LdcB, purified proteins were provided to collaborators in the Centre for Bacterial Cell Biology, Newcastle University, for the enzymatic activity to be tested against peptidoglycan isolated from $\Delta ldcB$ strains of *S. pneumoniae* and *B. subtilis* (**Section 4.3**). To assess whether LdcB is a zinc-dependent metalloprotease, the activity of *SpLdcB* was assessed in the presence of EDTA, zinc, copper and cobalt (**Section 5.5**). Finally, to determine the structural basis of the biochemical properties of the enzyme, as summarised above, crystal structures of LdcB with and without substrate or product mimics were sought with an overall goal of understanding the reaction mechanism used by these enzymes (**Sections 3.4, 4.2 & 5.4**).

Chapter 2:

Materials and methods

The following Chapter describes the materials and methods used in the overexpression, purification and biochemical characterisation of target proteins. It also describes the techniques used to crystallise and solve the structure of these proteins.

2.1 Molecular biology

2.1.1 Chemical suppliers

All chemicals, unless otherwise stated, were purchased from Sigma Aldrich (Poole) or Fisher Scientific (Loughborough). Molecular biology reagents and kits were purchased from Thermo Scientific (Reading) or Qiagen (Manchester).

2.1.2 General buffers and solutions

HEPES based buffers were made from HEPES free acid powder and adjusted to the correct pH at room temperature by the addition of sodium hydroxide. Tris-based buffers were made from Tris base powder and adjusted to the desired pH at room temperature with hydrochloric acid. The components of standard protein and DNA electrophoresis reagents and buffers are detailed in **Appendix 1**.

2.1.3 Growth media

All standard bacterial cultures were grown in Luria Bertani broth (LB; defined in **Appendix 1**) either in liquid form or on solid LB agar plates, which contained 1.5 % (w/v) agar prior to autoclaving. For proteins that were selenomethionine-labelled for structure determination, the corresponding methionine-auxotrophic strains of *Escherichia coli* were grown in a defined growth medium that is defined in **Appendix 1**. When required, kanamycin (50 µg/ml) or ampicillin (100 µg/ml) was added to growth media to select for plasmid transformants.

2.1.4 Strains and plasmids

A list of all bacterial strains used in this study, along with their genotypes, can be found in **Table 2.1**. A list of expression vectors used in this study can be found in **Table 2.2**.

Table 2.1: List of *E. coli* strains used in this study.

Strain	Description	Genotype
DH5α	Host strain, used for high efficiency cloning (Hanahan, 1983).	F- <i>endA1 hsdR17 (rk-mk+) supE44 thi-1recA1 gyrA (Nal^r) relA1Δ(lacZYAargF)U169(Φ80 lacZΔM15)</i>
B834 (DE3)	Parent strain of BL21, methionine auxotroph. Used for protein overexpression and labelling with Se-Met (Wood, 1966).	<i>E. coli B dcm ompT hsdS(rB-mB-) gal ,met, λDE3 phage</i>

Table 2.2: Plasmids and constructs used in this study

Plasmid	Description	Antibiotic selection	Source
pET28a	A high copy number plasmid with T7 promoter. A multiple cloning site exists with optional N- or C-terminal 6His•Tag. The N-terminal tag contains a thrombin cleavage site for tag removal following overexpression.	Kanamycin.	Dr. R. Cleverley ¹
pET28a- <i>SpLdcB</i>	Bacterial expression vector pET28a with N-terminal His•Tag and thrombin cleavage site fused to <i>SpLdcB</i> .	Kanamycin	Dr. R. Cleverley ¹
pET28a- <i>BsLdcB</i>	Bacterial expression vector pET28a with N-terminal His•Tag and thrombin cleavage site fused to <i>BsLdcB</i> .	Kanamycin	Dr. R. Cleverley ¹
p15TV-LIC- <i>BaLdcB</i>	Bacterial expression vector p15TV-LIC with N-terminal His•Tag and TEV cleavage site fused to <i>BaLdcB</i> .	Ampicillin	Dr. D. Meziane-Cherif ²
pET28a- <i>SpLdcB</i> -E204A	pET28a- <i>SpLdcB</i> with glutamate 204 alanine mutation.	Kanamycin	This study
pET28a- <i>SpLdcB</i> -E204H	pET28a- <i>SpLdcB</i> with glutamate 204 histidine mutation.	Kanamycin	This study
pET28a- <i>SpLdcB</i> -R120A	pET28a- <i>SpLdcB</i> with arginine 120 alanine mutation.	Kanamycin	This study

¹Newcastle Structural Biology Laboratory, Institute for Cell and Molecular Bioscience, Newcastle University, Newcastle upon Tyne, England

²Departement de Biologie Structurale et Chimie, Institut Pasteur, Unite de Bioinformatique Structurale, Paris, France

2.1.5 Preparation of competent cells

A single colony was selected from an LB agar plate and used to inoculate 100 ml LB. The culture was incubated at 37 °C with shaking at 200 rpm until an optical density at 600 nm (OD_{600}) was reached of 0.3 - 0.6. The culture was transferred to ice and left to cool for 30 minutes, prior to centrifuging at 4000 x *g* for 5 minutes at 4 °C. The pellet was resuspended in 25 ml ice-cold, sterile 0.1 M $MgCl_2$ and incubated on ice for 30 minutes. The culture was then centrifuged at 4000 x *g* for 5 minutes at 4 °C and resuspended in 25 ml ice-cold, sterile 0.1 M $CaCl_2$. Following incubation on ice for 30 minutes, the culture was again centrifuged at 4000 x *g* for 5 minutes at 4 °C and resuspended in ice-cold, sterile 0.1 M $CaCl_2$, 15 % (v/v) glycerol. 50 µl aliquots were taken and flash-frozen in liquid nitrogen. Competent cells were stored at -80 °C until required. One aliquot of each competent cell preparation was transformed with pET28a plasmid to assess transformation efficiency.

2.1.6 Transformation of competent cells

An aliquot of competent cells (50 µl) was defrosted on ice and 2 µl of plasmid (or 5 µl of mutagenesis reaction) was added. Cells were incubated with the vector for 30 minutes before heat-shocking at 42 °C for 45 seconds. Following 2 minutes of incubation on ice, 300 µl of LB media was added and the cells were allowed to recover for 1 hour at 37 °C before plating onto LB agar containing the appropriate antibiotic. A control transformation was performed in parallel, without the addition of plasmid, to ensure the cells did not contain any inherent antibiotic resistance. LB agar plates were incubated at 37 °C overnight.

2.1.7 QuikChange mutagenesis of *SpLdcB*

A full list of primers used in this study can be found in **Appendix 3**; pET28a-*SpLdcB* was used as the template in the production of all mutants. The mutagenic primers were phosphorylated at their 5' termini before deployment in the QuikChange reaction to promote the ligation and circularisation of the reaction products. The primer phosphorylation reactions contained 0.5 pmol/µl forward and reverse primer, 1 x T4 PNK reaction buffer, 1 mM ATP and 0.06 U/µl T4 polynucleotide kinase (Fermentas) in a final reaction volume of 20 µl. The reactions were incubated at 37 °C for 30 minutes and the T4 PNK inactivated by incubation at 70 °C for 5 minutes. The

phosphorylated primers were then used in QuikChange reactions, which contained 1 x *Pfu* polymerase reaction buffer, 50 ng dsDNA template, 5 µl of the phosphorylated primer mixture, 0.2 mM dNTPs and 2.5 U of *Pfu* polymerase (Promega). **Table 2.3** lists the QuikChange conditions used.

Table 2.3: Quickchange reaction conditions

Segment	Cycles	Temperature	Time
1	1	95°C	30 seconds
2	12–18	95°C	30 seconds
		55°C	1 minute
		68°C	12 minutes

2.1.8 Visualisation of mutagenic products by agarose gel electrophoresis

QuikChange reaction products were visualised by separation on 1 % (w/v) agarose gels stained with ethidium bromide (0.2 µg/ml) and 1 µl of 1 kb ladder (Fermentas) was used to estimate the sizes of the QuikChange products. All samples were mixed with 6x loading dye prior to loading into the gel. The gel was run at 150 V for 40 minutes before visualisation of the DNA on a UV transilluminator. **Figure 2.1** shows a representative agarose gel.

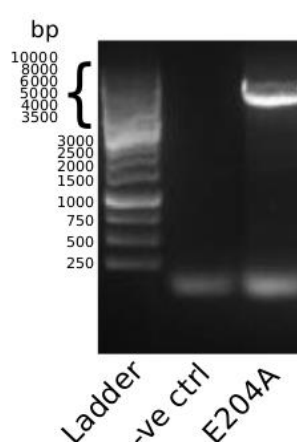


Figure 2.1: Representative mutagenesis agarose gel. DNA corresponding to a linearised pET28a-*SpLdcBE204A* can be seen in the E204A lane. The absence of DNA in the –ve control indicates successful digestion of the template plasmid by Dpn1.

2.1.9 Purification of plasmids from small-scale cultures

The products of successful QuikChange reactions were transformed into chemically competent DH5 α *E. coli* cells and single colonies were used to inoculate 10 ml LB containing the appropriate antibiotic. Cultures were allowed to grow overnight at 37 °C, with shaking at 200 rpm, prior to centrifugation at 3000 x *g* for 5 minutes at 4 °C. The plasmid was extracted from these pellets using GeneJET plasmid miniprep kit (Thermo Scientific) as per the manufacturer's instructions. Briefly, the pellet was resuspended in 250 μ l resuspension buffer and 250 μ l lysis buffer was added to lyse the cells. Following a 5 minute incubation in the lysis buffer, 350 μ l neutralisation solution was added to precipitate proteins and genomic DNA and the precipitant removed by centrifugation. The supernatant was added to a silica spin-column and centrifuged. The column was washed twice with 350 μ l ethanol wash buffer before elution of the bound plasmid with the addition of 50 μ l elution buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA).

2.1.10 DNA sequencing

All plasmids used in this study were sequenced by GATC biotech (Konstanz, Germany), using universal T7 primers (T7-minus1 and pET-RP) to obtain a forward and reverse read. BLAST (Altschul *et al.*, 1990) was used to compare the sequences of the cloned genes with those found in the KEGG database (Kanehisa *et al.*, 2004) or on the *Bacillus subtilis* specific database, Subtiwiki (Mäder *et al.*, 2012).

2.1.11 Analysis of proteins by SDS-PAGE electrophoresis

15% polyacrylamide gels were constructed using a Mini-PROTEAN® Tetra Cell (Biorad). The components of the stacking and resolving gels are listed in **Appendix 1**. The protein concentration of each sample to be electrophoresed was measured and 5-10 μ g of protein, and 10 μ l of prestained protein marker (Fermentas), was mixed with 6x SDS loading dye and heated to 95 °C for 2 minutes before loading. Gels were ran at 200 V for 45 minutes or until sufficient resolution had been achieved, as judged by the prestained marker. Gels were stained using Coomassie blue for 1 h before destaining overnight in destain solution.

2.2 Protein overexpression and purification

2.2.1 Large-scale overexpression and purification of target proteins (native and Se-Met substituted)

Recombinant *Bacillus subtilis* LdcB (*BsLdcB*; residues 61-273), *S. pneumoniae* LdcB (*SpLdcB*; residues 56-238) and *Bacillus anthracis* LdcB (*BaLdcB*; residues 55-243), were all overexpressed in the methionine auxotrophic B834 (DE3) strain of *E. coli*.

Overnight starter cultures were used to inoculate 1 l of LB containing the appropriate antibiotic. Cells were grown in LB media at 37 °C to an OD₆₀₀ of 0.4 and induced by the addition of IPTG to a final concentration of 1 mM. The culture was left to continue growing for 16 h at 18 °C for native proteins, or in minimal medium supplemented with selenomethionine for selenomethionine-labelled protein. Cells were harvested by centrifugation at 3000 x *g* for 25 minutes at 4 °C.

2.2.2 Purification of target proteins by metal affinity chromatography (Ni⁺-NTA)

The cell pellet was resuspended in 40 ml buffer A (50 mM HEPES/NaOH, pH 8.0, 300 mM NaCl, 20 mM imidazole) per litre of culture and lysed on ice for 3 min with a Sonopuls HD2070 (Bandelin) sonicator at 90% power, 80% cycle. The cell lysate was clarified at 33000 x *g* for 20 min and the supernatant loaded, using an AKTA prime (GE Healthcare), onto a 5 ml Ni⁺-NTA column (GE Healthcare) pre-equilibrated with buffer A. The column was washed with buffer A until the chromatogram returned to baseline before tightly-bound proteins were eluted with buffer B (50 mM HEPES/NaOH, pH 8.0, 300 mM NaCl, 500 mM imidazole) and the fractions containing the target protein were pooled. A typical chromatogram obtained from Ni⁺-NTA purification is shown in **Figure 2.2** (overleaf).

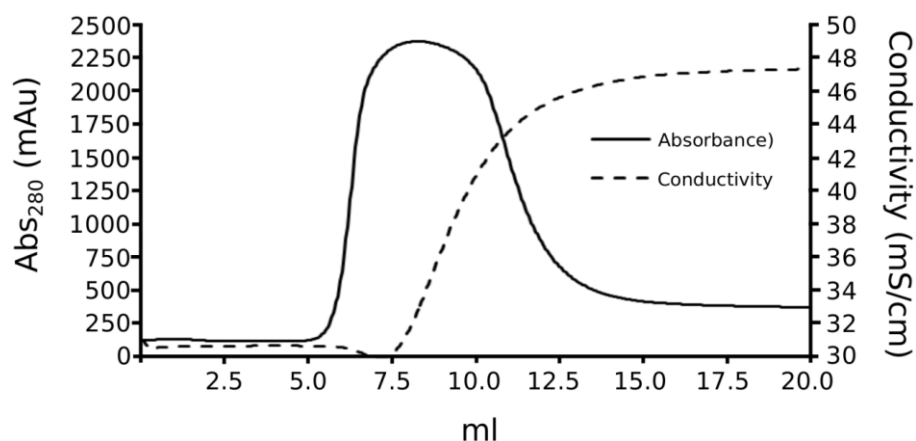


Figure 2.2: Representative chromatogram of *SpLdcB* elution from a Ni-NTA column.

2.2.3 Removal of hexahistidine tag by thrombin/TEV cleavage

Prior to size exclusion chromatography, the N-terminal hexahistidine tag was removed from *SpLdcB* and *BsLdcB* samples by digestion with thrombin (Sigma-Aldrich). A ratio of 1 unit of enzyme per mg of LdcB was added and the mixture incubated for 16 h at 20°C. Removal of the hexahistidine tag of *BaLdcB* was accomplished by incubating *BaLdcB* with a ratio of 1 µg TEV protease (provided by Dr. Vincent Rao, Newcastle University, UK) per 10 µg of *BaLdcB* at 4 °C for 16 h.

2.2.4 Purification of proteins by size exclusion chromatography (Superdex 75)

Each pooled protein was concentrated by ultrafiltration (Amicon Ultra-15 10K, Millipore) to a volume of 1 ml before loading, using an AKTApurifier (GE Healthcare), onto a Superdex S75 HiLoad 16/60 (GE Healthcare) gel filtration column pre-equilibrated in buffer C (10 mM HEPES/NaOH, pH 8.0, 100 mM NaCl). A typical size exclusion chromatogram can be observed in **Figure 2.3** (overleaf). The final purity of all protein preparations was judged by SDS-PAGE to be >98 % (**Figure 2.4**, overleaf).

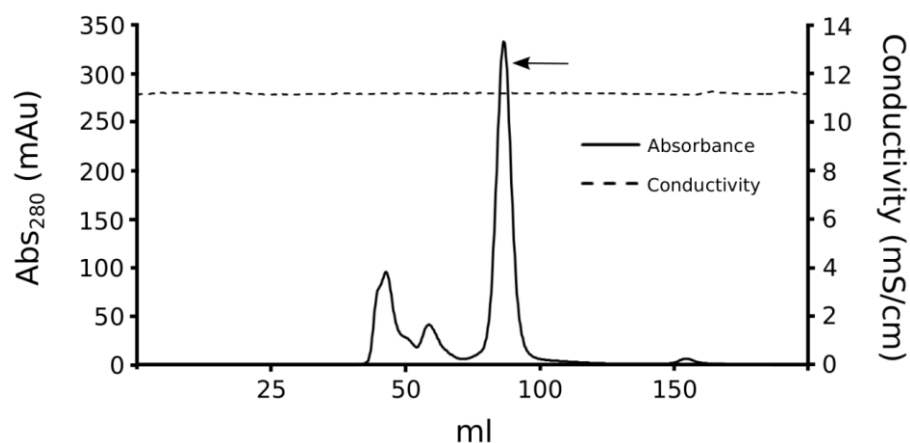


Figure 2.3: Representative chromatogram of *SpLdcB* elution from a Superdex S75 HiLoad 16/60 size exclusion column. The peak corresponding to *SpLdcB* is indicated by a black arrow.

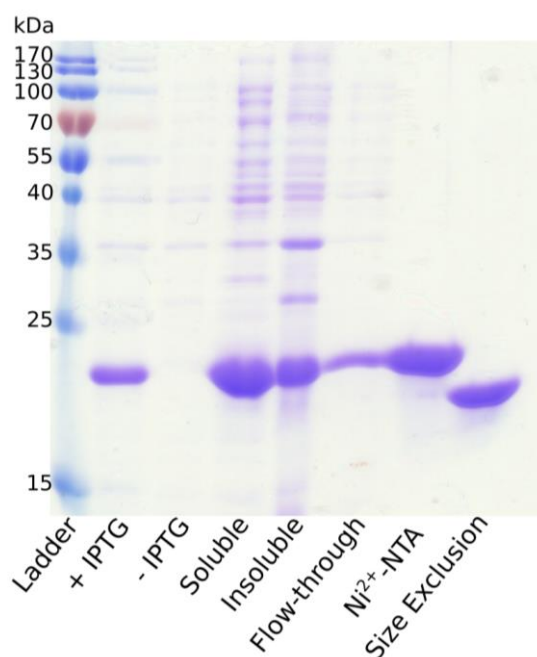


Figure 2.4: Representative purification gel. Overexpression of *SpLdcB* can be observed in the induced fraction (+ IPTG) but not in the uninduced (-IPTG). Furthermore the protein can be observed to be purified to homogeneity by Ni^{2+} -NTA purification. The shift in size observed from size exclusion is consistent with the loss of the N-terminal hexahistidine tag.

2.2.5 Extraction of native metals by dialysis against EDTA and subsequent buffer exchange

To remove any natively bound metal ions or those acquired during Ni^{2+} -NTA purification, LdcB proteins were dialysed against 50 mM EDTA in buffer C for 16 h at 4°C. Following dialysis, protein was concentrated to 250 μl and EDTA removed by

buffer exchange on a Superdex S75 HiLoad 10/40 (GE Healthcare) gel filtration column pre-equilibrated in buffer C. The removal of natively bound metals was confirmed by inductively coupled plasma mass spectrometry (ICP-MS).

2.3 Biophysical characterisation of purified proteins

2.3.1 Determination of protein concentration

Protein concentration was routinely determined by use of the Bradford assay, in which 15 μ l of protein sample was mixed with 200 μ l of Bradford reagent and made up to 1 ml with milliQ water. Absorbance readings were taken at 595 nm and protein concentration determined by comparison to a standard curve prepared from known concentrations of bovine serum albumin (BSA) (**Figure 2.5**). If the absorbance reading fell outside the values in the standard curve, the protein sample was diluted before the assay was performed.

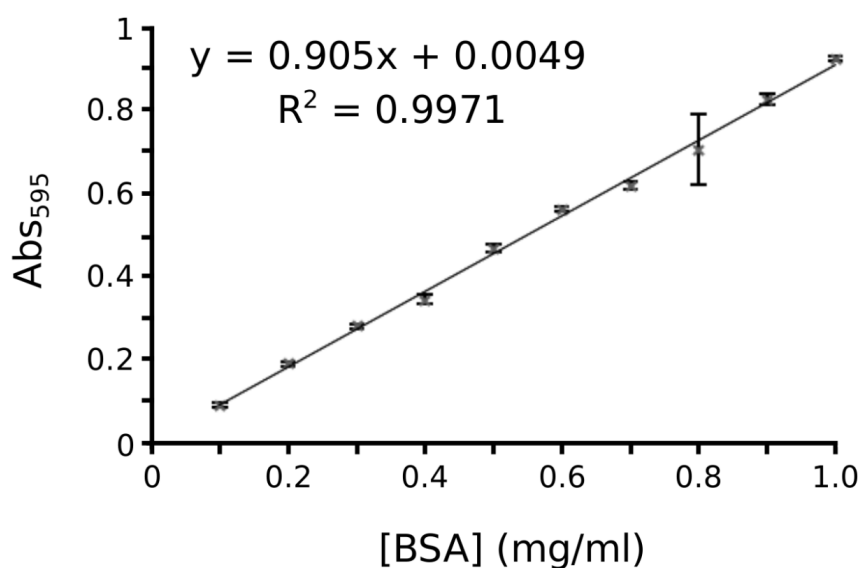


Figure 2.5: BSA concentration standard curve. The curve was produced by adding 15 μ l of BSA at known concentrations to 200 μ l Bradford reagent in a final volume of 1 ml. 3 repeats were performed at each data-point.

The Bradford assay is reliant on the biophysical properties of Coomassie Brilliant Blue G-250. Under acidic conditions the dye primarily absorbs light at λ_{max} 470 nm and appears red in colour (Compton and Jones, 1985). Upon interaction with the basic and aromatic residues found within the protein sample, a spectral shift occurs and the dye

absorbs light at λ_{\max} 610 nm and the dye appears blue. The greatest difference between the two absorbance maxima is at 595 nm, and therefore readings at this wavelength allow the most accurate quantification of the red:blue ratio (Compton and Jones, 1985). As the ratio of red:blue dye is due to the interaction of the dye with the amino acids present in the protein sample, the absorbance change upon protein binding is specific to that protein. To ensure the standard curve derived from BSA was appropriate to calculate the concentrations of *BsLdcB* and *SpLdcB*, the Beer-Lambert law was used to determine the concentrations of initial purifications. The Beer-Lambert law:

$$A = \epsilon cl$$

states that protein concentration (c) can be calculated if the extinction coefficient (ϵ), absorbance at 280 nm (A) and path length (l) are known. The extinction coefficient ($M^{-1}cm^{-1}$) was determined through the use of the ProtParam webserver: <http://web.expasy.org/protparam/> (Wilkins *et al.*, 1999). Readings obtained through the Bradford assay were found to closely match those obtained using absorbance at 280 nm.

2.3.2 Circular dichroism

Plane polarised light is made up of two components, left handed (rotating counter-clockwise) and right handed (rotating clockwise) (Kelly *et al.*, 2005). Under normal circumstances, both components have equal magnitude and sum to give radiation in one plane. However, should a material, such as a chiral chromophore, absorb more of one hand than the other, the sum of the remaining radiation becomes elliptical in nature (Kelly *et al.*, 2005).

The two chromophores usually found within proteins that can be exploited to obtain structural data are peptide bonds and aromatic side chains. Peptide bonds absorb below 240 nm (far UV) and so scans of between 180 – 250 nm give a general estimate of protein secondary structure. Aromatic side chains, absorbing between 260 and 320 nm (near UV), allow for the monitoring of protein tertiary structure, such as conformational changes induced by ligand binding (Kelly *et al.*, 2005).

Prior to obtaining CD spectra, the target protein was dialysed into 10 mM sodium phosphate buffer (pH 7.5), concentrated to ~0.1 mg/ml and placed in a 1 mm pathlength cuvette. Prior to collecting data, the CD spectropolarimeter was blanked against buffer. Scans were recorded from 180 – 250 nm on a JASCO J810 spectropolarimeter in 0.5 nm steps, with 10 repeats performed for all readings. All data were processed in CDSSTR software to also account for differences in protein concentration and plotted using Microsoft Excel.

2.4 Metal analysis of LdcB enzymes

2.4.1 Inductively coupled plasma mass spectrometry (ICP-MS)

Inductively coupled plasma mass spectrometry is a technique that can be used to identify and quantify the concentrations of metals and other elements typically found in biological samples, with a lower detection limit in the nM range (Ammann, 2007). Samples are dissolved in nitric acid prior to analysis to ensure the solubility of all metal ions. Following sample ionisation, magnetic fields are used to direct ions with specific mass/charge ratios to the detector, where the number of ions colliding with the detector is counted. ICP-MS can either be used to identify unknown metal cofactors found in protein samples or, if the metal is known, quantify the elemental concentrations (Ammann, 2007). Quantification is performed by comparing the number of ions colliding with the detector to a standard curve of the ion at known concentrations (Ammann, 2007).

All samples analysed by ICP-MS were diluted 10-fold in 2.5% (v/v) Suprapure HNO₃ (Merck) containing 20 ppb Ag⁺ and Pt⁺ ions prior to analysis (ICP-MS; Thermo Electron Corp., X-Series). ⁶⁷Zn, ⁶⁶Zn, ⁶⁵Cu, ⁶³Cu and ⁵⁹Co, ions were measured 100 times using the peak-jump method (30 ms dwell time, 5 channels, 0.02 AMU separation per isotope) and compared to matrix-matched elemental standards (VWR). Analyte counts were converted from ppb to µM and compared to the protein concentration, which was determined by the Bradford method.

2.4.2 Calculating binding affinities for *d*-block metals and BsLdcB

A common technique used to determine metal affinities is to compete the metalloprotein of interest against a metal chelator of known affinity for the metal ion.

The use of chelators that undergo a spectroscopic change upon the binding of a metal ion, such as 4-(2-pyridylazo)resorcinol (PAR), allows the competition assay to be monitored in an optical light spectrophotometer (Jefferson *et al.*, 1990).

PAR (200 μM) was incubated with ZnCl_2 (45 μM), CuCl_2 (25 μM) or CoCl_2 (10 μM) for 30 minutes to ensure the formation of 2:1 PAR:metal complexes. Increasing concentrations of *BsLdcB* (0 – 120 μM) were added to the PAR:metal complexes and the mixtures incubated at room temperature for 10 minutes to reach equilibrium. The absorbance of the mixture was monitored at 500 nm (Zn^{2+}) or 514 nm (Cu^{2+} and Co^{2+}) to allow for the calculation of the PAR-metal complex concentration using previously established extinction coefficients (McCall and Fierke, 2000). The concentration of the PAR-metal complex was plotted against the concentration of added protein and the following equation, adapted from (Badarau and Dennison, 2011), used to determine the affinity of *SpLdcB* for PAR:

$$K_d = \frac{[\text{BsLdcB}][\text{PAR}_2\text{M}^{2+}]}{\beta_{\text{PAR}}[\text{BsLdcB} \cdot \text{M}^{2+}][\text{PAR}]^2}$$

in which previously published formation constants (β_{PAR}) of PAR with zinc ($1.26 \times 10^{17} \text{ M}^{-2}$ (Zhou *et al.*, 1999)) or copper ($3.2 \times 10^{17} \text{ M}^{-2}$ (Turski *et al.*, 2012)) were used. M^{2+} represents the metal ion; Zn^{2+} , Co^{2+} or Cu^{2+} .

2.5 Biochemical characterisation of LdcB proteins

Unless otherwise stated, the biochemical characterisation of *SpLdcB* used protein samples that had been pre-incubated with 4 molar equivalents of ZnCl_2 .

2.5.1 Carboxypeptidase activity

2.5.1.1 Thin-layer chromatography (TLC)

End-point and timecourse TLC assays were performed to analyse the carboxypeptidase activity of LdcB. To assess the metal-dependent carboxypeptidase activity of LdcB, the tetrapeptide Ala-D-Gln-L-Lys-D-Ala (Tetra-D; Activotec) or Ala-D-Gln-L-Lys-L-Ala (Tetra-L; Activotec) at 20 mM were incubated with 700 μM of *SpLdcB* for 2 h at 37 °C in

a buffer of 10 mM HEPES/NaOH, pH 8.0, 100 mM NaCl in a total reaction volume of 10 μ l. A control sample of *SpLdcB* was incubated with 50 mM EDTA before reacting with the tetrapeptide. Control reactions also utilised lysozyme in place of *SpLdcB* to ensure the degradation of the tetrapeptide was specific to *SpLdcB* (data not shown). The reaction mixtures were spotted onto silica gel TLC aluminium foils (Sigma) and placed in a glass TLC tank containing the TLC running buffer of butan-1-ol, acetic acid, water in a 3:1:1 ratio, and the lid greased to seal the tank. The TLC foils were left for 40 minutes to achieve optimal species separation before staining with 0.1 % (w/v) ninhydrin solution in ethanol (Sigma). The TLC plates were dried and warmed with a domestic hairdryer to allow visualisation of the samples. To test *SpLdcB*, *BsLdcB* and *BaLdcB* (plasmid provided by Dr Djalal Meziane-Cherif, Institut Pasteur, France) dipeptidase and pentapeptidase activities, the above reaction conditions were used but with D-Ala-D-Ala dipeptide (Sigma) or Ala-D- γ -Gln-L-Lys-D-Ala-D-Ala (Sigma) substituting for the tetrapeptide. The *SpLdcB*-E204A, *SpLdcB*-E204H and *SpLdcB*-R120A (each at 700 μ M) were tested for activity against the Tetra-D substrate (20 mM) only, as described above.

Time-course analysis of *SpLdcB* followed a similar methodology as end-point TLC, with the following modifications. For *SpLdcB*-E204H and its comparison to both wild-type *SpLdcB* and *SpLdcB*-E204A, 20 mM tetrapeptide was incubated with 50 μ M protein in a total reaction volume of 30 μ l at 4 °C. Samples of the reactions (2 μ l) were taken every 15 minutes and the reaction stopped with the addition of ethanol to a final concentration of 50%. Following the spotting of the reaction mixtures onto silica gel TLC aluminium foils (Sigma), the running and staining of the TLC plates was performed as above. The activities of *SpLdcB*:copper and *SpLdcB*:cobalt were assessed and compared to *apo SpLdcB* (*SpLdcB*:zinc) by time-course, as described above.

2.5.1.2 Quantitative release of D-alanine

Purified *SpLdcB*, *BsLdcB* and *BaLdcB*, loaded with a molar equivalent of zinc, were flash-frozen in liquid nitrogen and provided to collaborators in the Institut Pasteur (Paris, France) for kinetic analysis. Carboxypeptidase activity was measured by Dr. Djalal Meziane-Cherif using the amino-acid oxidase-lactate dehydrogenase coupled assay (Badet *et al.*, 1984; Araoz *et al.*, 2000). Upon the release of D-alanine by LdcB, the amino acid is oxidised by D-amino acid oxidase (DAO) in an FAD-dependent

manner, producing pyruvate and releasing hydrogen peroxide. Pyruvate is then reduced by lactate dehydrogenase (LDH) in a NADH dependent manner, with the conversion of NADH to NAD⁺ followed at 340 nm to monitor the reaction (**Figure 2.6**). Catalase converts the hydrogen peroxide side-product into oxygen and water, permitting the cycling of FAD, a cofactor of DAO.

The reaction was performed in 0.1 ml containing 50 mM HEPES, pH 7.5, 0.2 mM NADH, 0.01 µg of lactate dehydrogenase (LDH), 0.03 µg of D-amino acid oxidase (containing 1 molar equivalent of FAD) (DAO), 100 units of catalase at 37°C, 15 µg of enzyme and various concentrations of L-Ala-D-iso-Gln-L-Lys-D-Ala (0 to 3 mM). The raw data provided by Djalal Meziane-Cherif was fitted using the Michaelis-Menten equation in sigmaplot v11.0 (Systat Software Inc.).

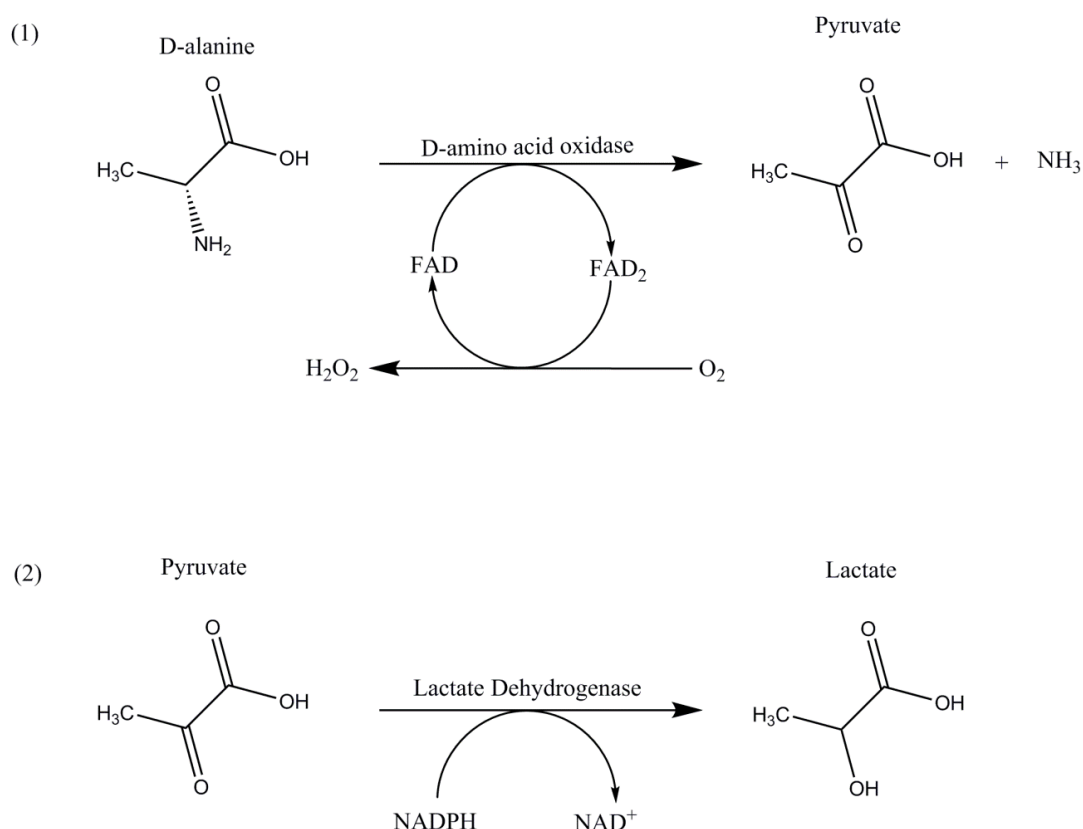


Figure 2.6: Reaction scheme for the conversion of D-alanine into lactate. (1) D-Alanine, liberated from the cell wall tetrapeptide by the activity of LdcB, is converted to pyruvate by D-amino acid oxidase, in an FAD dependent manner. (2) Pyruvate is reduced by lactate dehydrogenase to produce lactate. This reaction is NADPH dependent, with the conversion of NADPH > NAD⁺ allowing the reaction to be monitored at 340 nm.

2.5.2 Identification of natively bound ligands

The $2F_{\text{obs}} - F_{\text{calc}}$ electron density map of the solved *SpLdcB* structure indicated a small ligand was bound to the active site. Whilst the crystallisation condition did contain a mixture of amino acids, including the enzymatic product D-alanine, the density found in the active site appeared too large to be explained adequately by a single amino acid. To determine the identity of this unknown ligand, possibly a dipeptide that co-purified with *SpLdcB* from *E. coli*, two approaches were adopted. The first strategy compared the molecular masses of native vs denatured *SpLdcB*, as determined by electrospray mass spectrometry, whilst the second approach linked 1-fluoro-2-4-dinitrophenyl-5-L-alanine amide (FDAA or Marfey's reagent) to the unknown ligand to allow for its purification by reverse-phase HPLC and identification by comparison of elution volumes with control FDAA-amino acid derivatives.

2.5.2.1 Electrospray mass spectrometry

Purified *SpLdcB* was dialysed into 10 mM ammonium acetate buffer pH 8. Dialysis was performed for 24 hours, with buffer exchange occurring after 4 and 8 hours, to ensure trace ions would not interfere with the mass spectroscopy. Protein was then concentrated to 10 mg/ml (Amicon Ultra-15 10K, Millipore) and submitted for analysis by Dr Ault at the Mass Spectrometry facility, Astbury Centre for Structural Molecular Biology at the University of Leeds, United Kingdom.

2.5.2.2 Marfey's reagent and reverse-phase HPLC

FDAA can react with either L- or D- amino acids to produce diastereomers (**Figure 2.7**, overleaf). Diastereomers containing L-amino acids form stronger intramolecular bonds than those containing D-amino acids (B'Hymer *et al.*, 2003) and so experience a reduction in their polarity. This difference in polarity leads to D-derivatives having longer retention times on a reverse-phase HPLC column than their L-counterparts, and so HPLC can be used to purify a range of FDAA-amino acids from one another and from their own isomers (Bhushan and Brückner, 2004).

Purified *SpLdcB* (80 mg/ml) was denatured by heating and the insoluble material removed by centrifugal filtration (Amicon Ultra-15 10K, Millipore). The flow-through, theoretically containing the small molecules that were found in the active site of

SpLdcB prior to the heat-step, was then used as the substrate in a Marfey's reaction. An aliquot (200 μ l) of flow-through was transferred to an opaque glass reaction vial to which 150 μ l FDAA (50 mM final concentration) was added. The reaction was maintained at an alkaline pH by the addition of sodium bicarbonate (125 mM final concentration) and allowed to proceed for 1 hour at 37 °C with shaking. The reaction was stopped by the addition of HCl (50 mM final concentration) and the samples desiccated by freeze-drying for 48 hours. The samples were dissolved in 120 μ l 50 % (v/v) DMSO (dimethyl disulphate) prior to loading onto a C18 reverse-phase HPLC column pre-equilibrated with 20 mM ammonium acetate, pH 4. The samples were loaded onto a 1260 Infinity Quaternary LC System HPLC and FDAA-derivatised amino acids eluted by a gradient of acetonitrile (0-50%). Retention times were compared with those produced by control reactions containing D- or L-alanine.

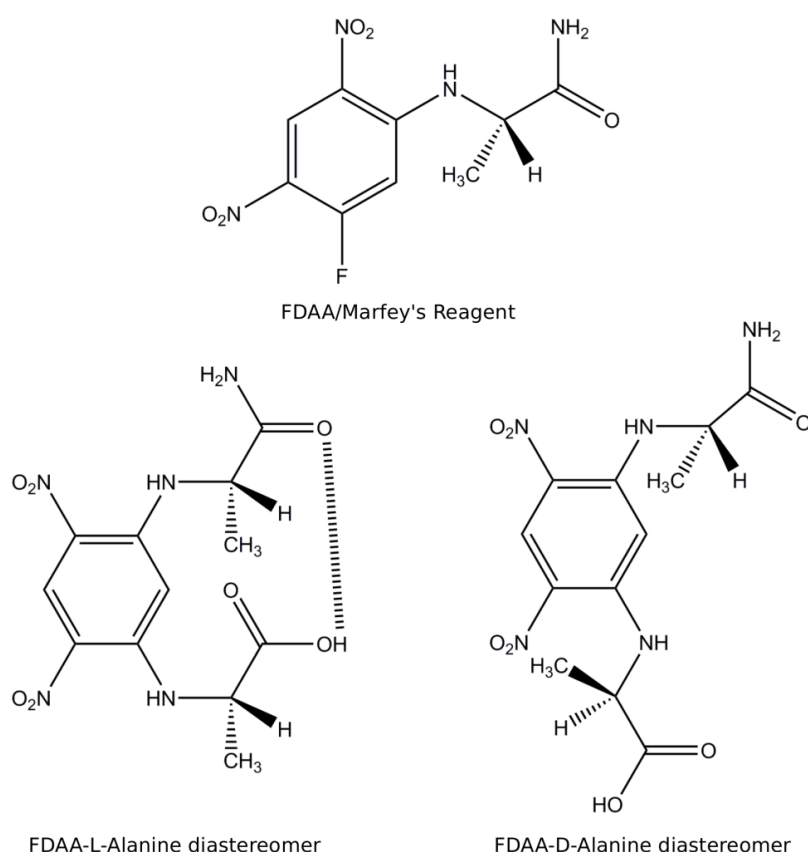


Figure 2.7: Marfey's reagent derivatisation of L and D alanine. The L-alanine diastereomer forms stronger intramolecular contacts than the D-disteromer. Hydrogen bonds are shown as a dashed line. Figure adapted from B'Hymer *et al.*, (2003).

2.6 Crystallisation of target proteins

2.6.1 Crystallisation trials

Both *BsLdcB* and Se-Met substituted *SpLdcB* were concentrated to 15 mg/ml by ultrafiltration (Amicon Ultra-15 10K, Millipore) and four molar equivalents of zinc chloride were added. Crystallisation conditions at 20 °C for both proteins were obtained by sparse-matrix screening using the commercial screens PACT, JCSG+ and Morpheus (Molecular Dimensions) in 96-well MRC crystallisation plates (Molecular Dimensions). A Mosquito (TTP Labtech) was used for all initial screens, setting up two sitting drops (100 nl + 100 nl and 200 nl + 100 nl of protein and crystallisation condition, respectively) per condition. The best *SpLdcB* crystals were obtained after 48 hours from Morpheus screen condition H8, containing a 0.1 M mixture of amino acids (glycine, L-glutamate, and racemates of alanine, serine and lysine), 0.1 M buffer system 2, pH 7.5 and 37.5 % MPD/PEG1000/PEG3350. By contrast, a single hit was achieved for *BsLdcB*, with several plates growing from a single nucleation point in 0.1 M phosphate/citrate buffer, pH 4.2, 40 % PEG 300 over a period of ~2 months. Crystals of native *SpLdcB* with bound MurNAc-L-Ala-D-iso-Gln-L-Lys-(D-Asn) (Invivogen) were obtained as above, with the additional step of incubating the protein with the ligand at 2 mM final concentration for 30 min at room temperature prior to crystallisation screening. Crystallisation of metal-substituted *SpLdcB* occurred as above, with Cu²⁺ or Co²⁺ being added once zinc was removed. Successful crystallisation solutions were optimised manually in hanging drops in 24-well Linbro plates (Molecular Dimensions).

2.6.2 In-house screening of crystals for diffraction

Crystals were harvested in a suitably sized cryoloops (Hampton) and cryo-protected in paratone-N oil before being flash cooled in liquid nitrogen (-173 °C). Crystals were mounted on a copper anode Rigaku MSC microfocus generator and data collected at a 1.54 Å on a Raxis IV⁺⁺ image plate detector. Two images per crystal were collected, at relative phi angles of 0 ° and 90 °, to allow for indexing in MOSFLM (Leslie, 2006). Crystals were ranked based upon the diffraction pattern and resolution, with the best being collected at the Diamond Light Source (DLS) synchrotron.

2.6.3 Data collection

2.6.3.1 Data collection for molecular replacement

Diffraction data for *BsLdcB* were collected at DLS on beamline IO2 on 28/9/12 at a wavelength of 0.97 Å.

Diffraction data of native *SpLdcB* with bound MurNAc-L-Ala-D-iso-Gln-L-Lys-(D-Asn) were collected at DLS on beamline IO2 on 11/03/13 at a wavelength of 0.92 Å.

2.6.3.2 SAD data collection

Data collection for Se-Met substituted *SpLdcB* took place on 28/9/12 on beamline IO2 at DLS. A fluorescence scan was performed to determine the optimal wavelength for SAD data collection. Data were collected at a wavelength of 0.98 Å, sufficiently close to the selenomethionine *K* atomic absorption edge to permit phasing by the single wavelength anomalous dispersion method.

2.6.3.3 Data collection for anomalous difference maps

Data collection for *SpLdcB* with bound Co^{2+} occurred on 10/2/13 on beamline IO4 at DLS. Data were collected at 1.605 Å, above the *K* edge of Co^{2+} but below that of Cu^{2+} and Zn^{2+} , and at 1.631 Å, below the *K* edge of all three metals. The presence of anomalous signal in the dataset above the *K* edge of Co^{2+} and its absence in the dataset collected below the edge allowed for the calculation of anomalous difference maps.

Data collection for *SpLdcB* with bound Cu^{2+} occurred on 10/2/13 on beamline IO2 at DLS. Data were collected at 1.494 Å, above the *K* edge of Cu^{2+} and Co^{2+} but below that of Zn^{2+} , and at 1.631 Å, below the *K* edge of all three metals. The presence of anomalous signal in the dataset above the *K* edge of Cu^{2+} and its absence in the dataset collected below the edge allowed for the calculation of anomalous difference maps.

2.6.4 Data processing

XDS (Kabsch, 2010) was used to index and scale all *SpLdcB* and *BsLdcB* diffraction data. During the indexing step, the reflections present on the images are identified and assigned a Miller index (*hkl*). In addition, the crystal geometry is determined to allow for the integration of the reflections. The data is then scaled to average all symmetry

equivalent reflections (i.e. Friedel pairs h,k,l and $-h,-k,-l$) and sum partially recorded reflections into one, normalised set of structure factors, I_{hkl} .

During scaling, the quality of data is assessed and an R-, or reliability, factor is produced. This value, known as R_{merge} if data are being merged between datasets, or R_{sym} if a single crystal is used, is the extent at which the observed intensities of each unique reflection deviates from the one another. R_{sym} can be calculated using the following formula:

$$R_{\text{sym}} = \frac{\sum_{hkl} \sum_j |I_{hkl,j} - (I_{hkl})|}{\sum_{hkl} \sum_j I_{hkl,j}}$$

where $I_{hkl,j}$ is the intensity of a reflection, j , and (I_{hkl}) is the mean intensity for all reflections within the averaged unique reflection. A small R_{sym} indicates only small differences between equivalent reflections. Unfortunately, whilst large multiplicity within a dataset should increase the quality of the data, as the accuracy to which intensities are can be averaged is improving, the R_{sym} value often increases. For datasets with large multiplicities, it is recommended that $R_{\text{p.i.m}}$ be reported in place of R_{sym} (Weiss, 2001), as this accounts for the large multiplicity of the dataset by the addition of an $(1/N-1)$ term to the R_{sym} equation:

$$R_{\text{p.i.m}} = \frac{\sum_{hkl} \frac{\sqrt{1}}{N-1} \sum_{j=1}^n |I_{hkl,j} - (I_{hkl})|}{\sum_{hkl} \sum_j I_{hkl,j}}$$

The $(N/N-1)$ term ensures that the $R_{\text{p.i.m}}$ value decreases with increased multiplicity and so reflects the increasingly accurate averaged intensity of the unique reflection.

An additional metric of data quality is the $I/\sigma I$ value, which is the ratio between the averaged intensity and the standard deviation of the averaged intensity of all unique reflections. Higher $I/\sigma I$ values are therefore indicative of a carefully collected dataset.

2.6.5 Determination of an initial model by molecular replacement

To calculate an electron density map from data collected by X-ray crystallography we require not only the intensities of each recorded reflection, but also the phase of the photons that contribute to the reflection. Whilst the intensities are collected, the phases are lost during data collection (Grunbaum, 1975).

Molecular replacement uses the phases calculated from a similar structure to estimate those for the unknown structure and then refines the estimated phases until they approximate the true phases (Scapin, 2013). Although the structure of the molecule for which the phases are required is unknown, sequence identity can be used to suggest proteins that may be similar, with proteins sharing sequence identities of ~30% usually adequate to use for molecular replacement.

Before the phases of the unknown structure can be estimated, the known structure must be orientated and positioned correctly in the asymmetric unit. Two methods are generally used for this, Patterson-based (e.g. MolRep) and likelihood-based (e.g. PHASER).

Patterson-based molecular replacement calculates Patterson maps for both structures. These maps contain peaks that represent the vectors between atoms, which can be either intermolecular or intramolecular (Turkenburg and Dodson, 1996). Intramolecular vectors are dependent only on the orientation of the molecule, not on its position and so similar intramolecular vectors can be observed between two structures if they are orientated in the same manner. Once the two molecules are correctly orientated, the intermolecular vectors, which account for both orientation and position, can be used to position the molecules (Turkenburg and Dodson, 1996).

Likelihood-based molecular replacement tests the orientation and position of the model for consistency with the data, with the ideal solution being the model that best fits the observed data (McCoy *et al.*, 2007). Like the Patterson-based method, the rotation and translations functions are calculated independently.

PHASER (McCoy *et al.*, 2007) was used to solve the structures of *BsLdcB*, *SpLdcB* with bound ligand, *SpLdcB*-Co²⁺ and *SpLdcB*-Cu²⁺ and is a part of the CCP4 suite (Collaborative Computational Project, 1994). Potential solutions were monitored by

observing the ‘contrast’ score in MolRep or ‘LLG’ score in PHASER before one round of rigid body refinement in Refmac5 (Murshudov *et al.*, 1997) and manual inspection in COOT (Emsley and Cowtan, 2004). **Table 2.4** contains a list of structures solved by molecular replacement, in addition to the models used to solve them.

Table 2.4: A list of all structures solved by molecular replacement, the models used and their sequence similarity

Structure	Model	Sequence identity (%)
<i>BsLdcB</i>	<i>SpLdcB</i> – this study	30
<i>SpLdcB</i> with bound ligand	<i>SpLdcB</i> – this study	100
<i>SpLdcB</i> -Co ²⁺	<i>SpLdcB</i> – this study	100
<i>SpLdcB</i> -Cu ²⁺	<i>SpLdcB</i> – this study	100

2.6.6 Heavy atom substructure solution (SAD)

The intensity of scattering produced by an atom during data collection is known as the atomic form factor ($f^0(\theta)$) (Dauter *et al.*, 2002). However, when a heavy atom is introduced as a scatterer, the atomic form factor must be modified to account for an additional signal produced by these atoms at specific wavelengths; at these wavelengths, the incident X-ray beam provides enough energy for an electron of a low energy shell to move to a higher energy shell and in doing so, emit an anomalous signal (Dodson, 2003). The full structure form equation is therefore:

$$f(\theta, \lambda) = f^0(\theta) + f'(\lambda) + if''(\lambda)$$

where $f'(\lambda) + if''(\lambda)$ represent the anomalous scattering of these heavy atoms at particular wavelengths. In most proteins, there are no anomalous scatters, and so the contributions of f'' and f' are zero. In this case, Friedel’s Law holds true and symmetry related reflections will possess identical intensities but have opposite, but complementary, phases ($h,k,l = -h, -k, -l$). When an anomalous scatterer is present, the contributions of if'' to the atomic form factor cause a breakdown in Friedel’s Law resulting in symmetry related reflections having different intensities and non-

complementary phases. The differences in the atomic form factors can be used to determine the phases (Dauter *et al.*, 2002; Dodson, 2003).

The use of a single wavelength anomalous dispersion results in two possible phases, of which only one is correct. This ambiguity is usually solved by examining the electron-density map that is produced for interpretable features such as alpha-helices (Dauter *et al.*, 2002).

The structure of *SpLdcB* was solved using SHELX (Sheldrick, 2008) as implemented in the HKL2MAP interface (Pape and Schneider, 2004), not only to locate the selenium heavy atom substructure, but also to determine the correct anomalous hand and to apply solvent flattening. An initial model was built by PHENIX.AUTOSOLVE (Adams *et al.*, 2011), which successfully positioned 180 out of 185 amino acids.

2.6.7 Calculation of anomalous difference maps

In addition to allowing phasing, anomalous signals can be used to calculate an anomalous difference map, confirming the presence, or absence, of a heavy-atom scatterer. Phases were obtained by molecular replacement, using *SpLdcB* as a search model, and combined with anomalous data using CAD, of the CCP4 software suite (Collaborative Computational Project, 1994). FFT (Ten Eyck, 1985; Read and Schierbeek, 1988) was then used to calculate the anomalous difference map.

2.6.8 Model building and refinement

All electron density maps were improved by manual building cycles in COOT (Emsley and Cowtan, 2004), interspaced with restrained refinement either Refmac5 (Murshudov *et al.*, 1997) or PHENIX.REFINE (Afonine *et al.*, 2012). The reliability of the model was assessed by comparing the measured amplitudes with those calculated from the model, producing an R-factor:

$$\frac{\sum_h ||F_{obs}| - |F_{calc}||}{\sum_h |F_{obs}|}$$

where h is the sum of all reflections, F_{obs} are the observed amplitudes and F_{calc} are the calculated amplitudes. As the measured amplitudes were used as data during model building, bias could be introduced that artificially lowered the R-factor. This bias can be avoided by calculating an R-factor, termed the R_{free} , from a subset of reflections (usually 5%) that were not used in refinement. Once convergence was reached, water molecules and other ligands or small molecules were added.

2.6.9 Model validation

Models were validated through the use of COOT (Emsley and Cowtan, 2004), MOLPROBITY (Chen *et al.*, 2010) and RAMPAGE (Lovell *et al.*, 2003) to check geometry and Ramachandran outliers in addition to clashes.

Chapter 3:

Structural characterisation of LdcB from *S. pneumoniae* and *B. subtilis*

3.1 Introduction

In most species the terminal D-alanine of the pentapeptide stem is hydrolysed by DD-carboxypeptidases from the class C penicillin binding protein (PBP) family (Lovering *et al.*, 2012). LD-carboxypeptidases subsequently hydrolyse the L-Lys[A2pm]—D-Ala bond to release the tripeptide. Though tripeptides are found in many bacteria, very few of the corresponding LD-carboxypeptidases have been identified. This Chapter concerns one such LD-carboxypeptidase, called LdcB, the genetic disruption of which causes severe morphological defects (Barendt *et al.*, 2011). Mutant strains develop into small, spherical cells or elongate into long, thick cells, as a result of division asymmetry produced by random septum localisation (Barendt *et al.*, 2011). The morphological defects observed upon deletion of *ldcB* suggest that it could be of interest as an antibiotic target, with the design of these antibiotics aided by a crystal structure of LdcB. In this chapter we describe the crystallisation and structure determination of LdcB from *S. pneumoniae* and *B. subtilis* (**Sections 3.2 and 3.3**), and compare the structures with each other and other, related enzymes (**Section 3.4**).

3.2 Crystallisation trials

3.2.1 Crystallisation trials of native *SpLdcB* and *BsLdcB*

Native *SpLdcB* and *BsLdcB* were purified as described previously (**Section 2.2**). Both proteins were found to be >95 % pure, as determined by SDS-PAGE (**Section 2.2**) following a 3-step purification process. The proteins were each concentrated to 15 mg/ml for crystallisation trials, which were conducted in 96-well sitting-drop vapour diffusion trays utilising a Mosquito (TTP labtech) crystallisation robot and a range of commercially-available screens (Molecular Dimensions). Crystals grew in a number of conditions over a period of 7 days for both *SpLdcB* and *BsLdcB*. Rod-like *SpLdcB* crystals were observed to grow from JSCG+ (Newman *et al.*, 2005) solution F8 (2.1 M malic acid) (**Figure 3.1**) whilst *BsLdcB* grew from PACT premier (Newman *et al.*, 2005) solution E9 (0.2 M potassium sodium tartrate 20% PEG 3350) in clumps of needles originating from a single nucleation point (**Figure 3.1**, overleaf).

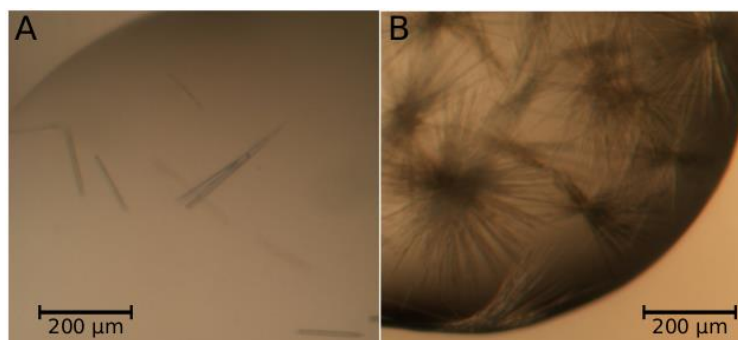


Figure 3.1: Initial crystal forms of *SpLdcB* (A) and *BsLdcB* (B).

Despite repeated attempts at optimising these crystallisation conditions, *BsLdcB* crystals failed to diffract and the best *SpLdcB* crystals diffracted to a maximum resolution of 3.6 Å on the microfocus beamline I24 at the Diamond Light Source. Additionally, the lack of a suitable molecular replacement model in the PDB at the time of data collection prevented us from using these data for structure solution. The data collection statistics can be found in **Table 3.1**.

Table 3.1: Summary of X-ray data collection statistics for initial *SpLdcB* crystals.

	<i>SpLdcB</i>
Data collection	
Space Group	<i>P</i> 21 21 2
Unit cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	33.8, 78.3, 149.9
α , β , γ (°)	90.00, 90.00, 90.00
Wavelength (Å)	1.28
Resolution (Å)	78.3-3.60
	(3.94-3.60)*
Multiplicity	12.4 (12.1)
R_{merge}^a	0.194 (0.842)
$I/\sigma I$	11.7 (2.3)
Completeness (%)	96.8 (85.4)

^a $R_{\text{merge}} = \sum_h \sum_i |I_i(h) - \langle I(h) \rangle| / \sum_h \sum_i I_i(h)$, where $I_i(h)$ and $\langle I(h) \rangle$ are the *i*th and mean measurement of the intensity of reflection *h*.

3.3 Crystallisation of Se-Met substituted *SpLdcB* and the addition of zinc

3.3.1 Incorporation of selenomethionine

The lack of a suitable molecular replacement model required to solve the structure of *SpLdcB* meant that an alternative phasing method was required. The substitution of

methionine by selenomethionine is a common method of introducing heavy atoms into protein crystals, from which single anomalous dispersion (SAD) data can be collected and used to solve the phase problem (Doublié, 1997). *E. coli* B834 cells containing pET28a-*SpLdcB* were grown in Se-Met media (**Appendix 1**) and the overexpression induced by the addition of 1 mM IPTG. A control culture was grown alongside, to which selenomethionine was not added. The control culture ceased growth after approximately 30 minutes, whilst the culture with selenomethionine continued growing, indicating selenomethionine incorporation was successful. Se-Met labelled *SpLdcB* was purified as previously described (**Section 2.2**) and used in crystallisation trials against commercially available crystallisation conditions (Molecular Dimensions) in 96-well sitting-drop vapour diffusion trays. Unfortunately, these crystals were of similar morphology to those produced by native *SpLdcB* and diffracted poorly in-house.

3.2.2 Addition of zinc and effect on crystallisation of LdcB

SpLdcB shares sequence identities of 38%, 36% and 30% to the known zinc metallo-proteins VanY, VanXYg and VanXYc (Binda *et al.*, 2012), indicating that LdcB is a metallo-enzyme; unfortunately, although the structures of VanXYg and VanXYc have subsequently been solved, they were not available to use as molecular replacement models during structure solution. Moreover, the amino acids that co-ordinate the zinc ion in the structures of VanX (H116, D123, H184) (PBDId: 1R44), VanXYg (H107, D114, H167) (PBDId: 4f78) and VanXYc (H95, D102, H156) (PBDId: 4MUT) are all conserved in LdcB. Consequently, the metal content of the Se-Met labelled *SpLdcB* was assessed by inductively coupled plasma mass spectrometry (ICP-MS) (Thermo X-series). Following purification, *SpLdcB* was diluted to 5 mg/ml in 2.5 % (v/v) Suprapure HNO₃ to denature the protein and ensure the solubility of the liberated metal ions. The sample was then ionised and a quadrupole magnet used to direct ions of specific mass/charge ratios towards a detector where the number of collisions were recorded. A standard curve of known metal concentrations was used to convert the number of 'hits' to μ M concentrations for comparison with protein concentration. *SpLdcB* purifies with roughly 50% zinc occupancy in addition some nickel, presumable as a consequence of Ni⁺-NTA purification (**Figure 3.2**, overleaf). The addition of excess zinc, followed by desalting of the proteins (HiTrap Desalting, 5 ml [GE Healthcare]) allowed the full occupancy of the

protein with zinc at an almost 1:1 stoichiometry, with the large excess of free zinc indicating a single, fully occupied binding site (**Figure 3.3**).

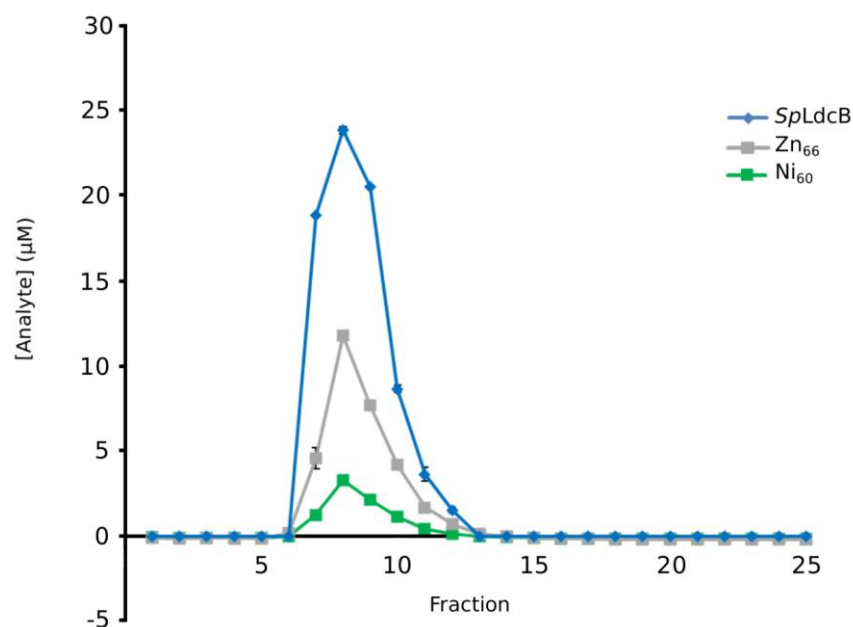


Figure 3.2: Metal content analysis of purified *SpLdcB*. The concentration of zinc and nickel of *SpLdcB* after purification by Ni-NTA and size exclusion chromatography was assessed by ICP-MS. Protein concentrations were determined by the Bradford method. All readings were taken in triplicate and standard deviation shown as error bars.

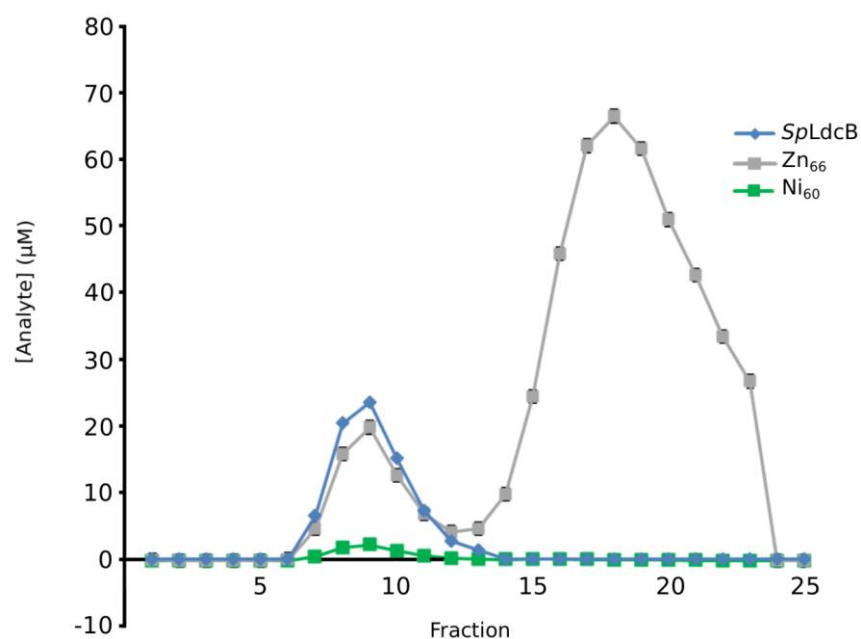


Figure 3.3: Metal content analysis of *SpLdcB* following zinc addition. The concentration of zinc and nickel of *SpLdcB* after the addition of zinc to purified protein, followed by separation on a HiTrap Desalting, 5 ml column was assessed by ICP-MS. Protein concentrations were determined by the Bradford method. All readings were taken in triplicate and standard deviation shown as error bars.

The presence of a ligand or a cofactor can often stabilise proteins, and stable proteins are more likely to form crystals. Thus, a 4:1 excess of zinc chloride was added to *SpLdcB* prior to screening for crystallisation. Crystallisation experiments were conducted as described above (**Section 3.2.1**). The addition of zinc chloride had a marked effect on the crystals of *SpLdcB*, producing 3D crystals in the Morpheus (Gorrec, 2009) H8 condition (0.1 M mixture of amino acids, 0.1 M buffer system 2, 37.5% MPD/PEG1000/PEG3350) within 48 hours of the experiment. The crystals of *BsLdcB* were not improved by the addition of excess zinc.

The zinc-loaded, Se-Met labelled crystals were optimised in 24 well hanging-drop vapour diffusion trays as described previously (**Section 2.6**) using solutions obtained from Molecular Dimensions. The unique combinations of precipitants, buffer systems and additives found in the Morpheus screen make optimisation a non-trivial endeavour. The buffer system 2 (0.1 M sodium HEPES, MOPS (acid) pH 7.5) concentration was kept constant at 0.1 M, while the concentration of amino acids (L-Na-Glu, Ala (racemic), Gly, Lys (racemic), Ser (racemic)) and precipitants (MPD/PEG1000/PEG3350) were varied from 0 to 0.2 M in 0.05 M increments and from 15% to 52.5% in 7.5% increments, respectively. The crystals used for data collection grew from the original crystallisation condition (0.1 M mixture of amino acids, 0.1 M buffer system 2, 37.5% MPD/PEG1000/PEG3350) at 21°C with a protein:mother-liquor ratio of 2:1. The final dimensions of the optimised crystals were $\sim 0.1 \times 0.1 \times 0.15 \text{ mm}^3$ (Figure 3.4).

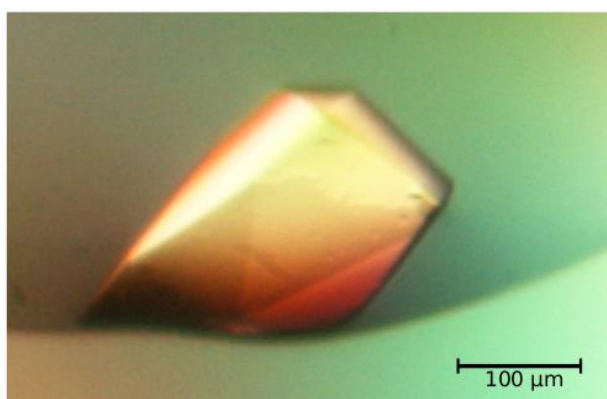


Figure 3.4: Improved crystal form of *SpLdcB*. The 3D-crystal form was obtained upon the addition of excess zinc prior to crystallisation.

Several Se-Met *SpLdcB* crystals were harvested and cryo-protected in paratone-N oil before flash-cooling in liquid nitrogen. Diffraction was assessed in-house, with two images being taken at relative phi angles 90° apart. Crystals were ranked based upon spot sharpness, mosaicity and strength of diffraction. The best crystals were taken to Diamond Light Source for data collection on beamline IO2. Fluorescence scans were performed on the *SpLdcB* crystals to confirm the incorporation of selenomethionine (**Figure 3.5**). The scans also identified the appropriate wavelength at which the single anomalous dispersion (SAD) data for selenium should be collected.

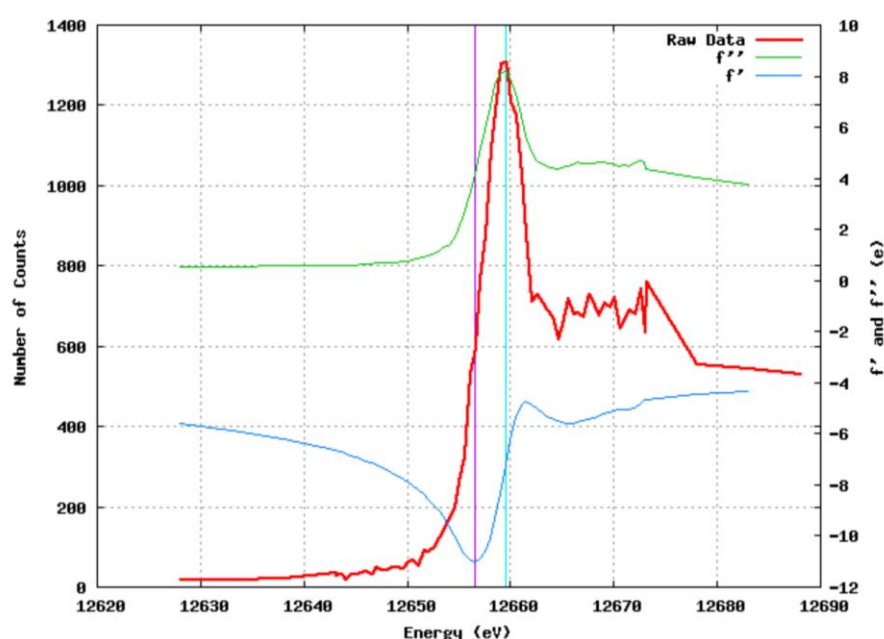


Figure 3.5: Metal content analysis (MCA) fluorescence scan of *SpLdcB* crystals. MCA scan assessing the incorporation of selenium into the crystals. The peak (red line) corresponds to the K-edge excitation wavelength of selenium. The blue line (f') corresponds to the real component of anomalous scattering around the K-edge of selenium, while the green line (f'') corresponds to the imaginary component. The light blue vertical line indicates the position of the peak (maximum of f''), whilst the purple vertical line indicated the position of the inflection (minimum of f').

A dataset with a high multiplicity is desirable for solving a structure by SAD, as sampling every reflection multiple times increases the accuracy and precision at which small anomalous differences can be measured (Weiss, 2001). A total of 7200 images were collected with 0.1° oscillations and 0.1 s exposure per image at a wavelength of 0.979 Å, the wavelength at which the fluorescence scan (**Figure 3.5**) indicated would produce the greatest amount of anomalous scatter from the selenium heavy atom substructure. An example of the diffraction obtained is shown in **Figure 3.6**.

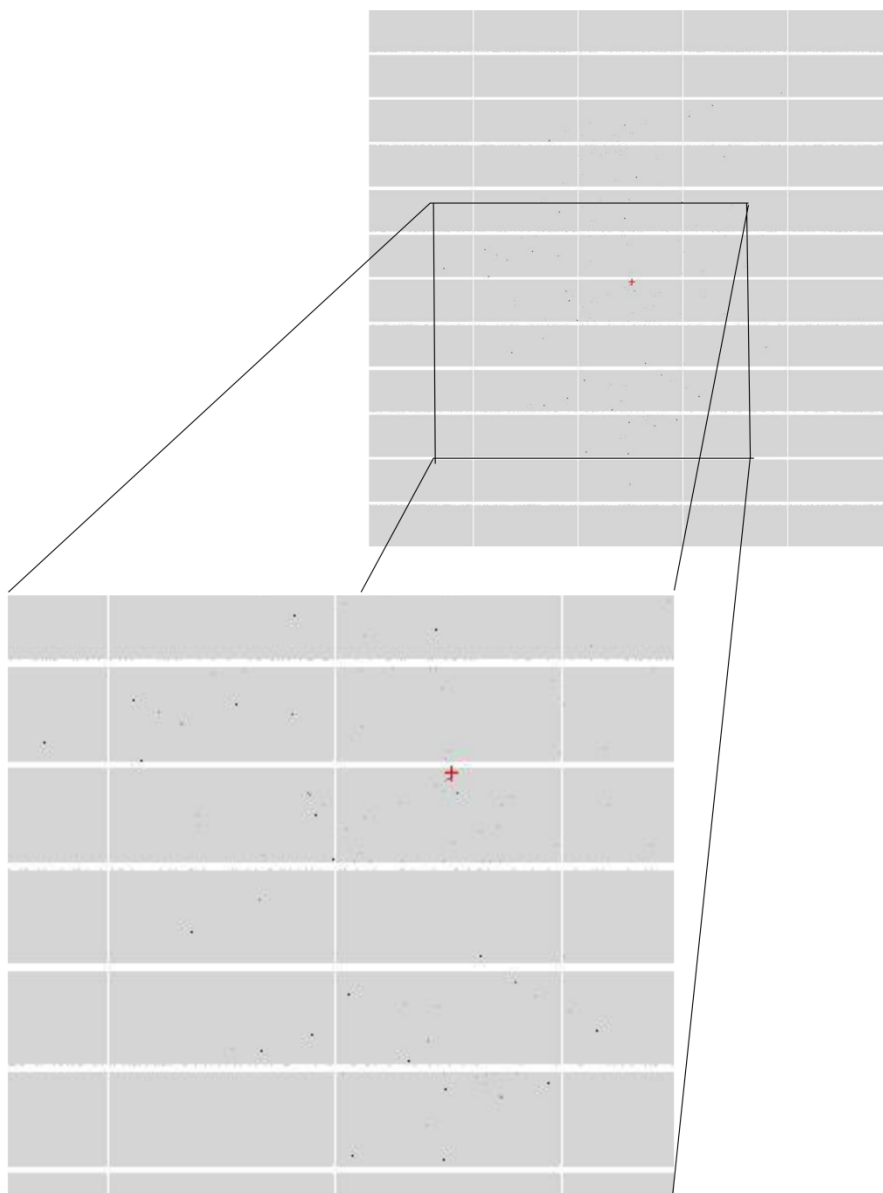


Figure 3.6: An example diffraction pattern produced by Se-Met labelled *SpLdcB* crystals. The crystal diffracted to 1.8 Å. The spots are difficult to define with the human eye due to the resolution of the Pilatus detector and the use of fine slicing in data collection.

3.3.3 Data processing and structure solution of Se-Met substituted *SpLdcB*

Diffraction data were integrated with XDS (Kabsch, 2010) and reduced with SCALA (Evans, 2005). The crystals were assigned to the spacegroup *I*222 with unit cell dimensions of $a = 48.8$, $b = 60.4$, $c = 138.2$ Å, based upon the systematic absence of reflections where $h+k+l=2n+1$ and scaling statistics. Data collection statistics are summarised in **Table 3.2**. The calculated Matthews' coefficient (Matthews, 1968) of

2.4 (Å³/Da), suggested that one protein molecule was present within the asymmetric unit with a solvent content of 49.0 %.

Table 3.2: Summary of X-ray data collection statistics for *SpLdcB*.

PDBid	<i>SpLdcB</i> 4OX5
Data collection	
Space Group	<i>I</i> 222
Unit cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	48.8, 60.4, 138.2
α , β , γ (°)	90, 90, 90
Wavelength (Å)	0.98
Resolution (Å)	20.0-1.80 (1.84-1.80)*
Multiplicity	24.3 (14.2)
Anomalous multiplicity	12.5 (6.9)
R_{merge}^a	0.077 (0.108)
<i>I</i> / σ <i>I</i>	34.4 (2.4)
Completeness (%)	98.6 (86.9)
Anomalous completeness	98.2 (83.9)

^a $R_{\text{merge}} = \sum_h \sum_i |I_i(h) - \langle I(h) \rangle| / \sum_h \sum_i I_i(h)$, where $I_i(h)$ and $\langle I(h) \rangle$ are the *i*th and mean measurement of the intensity of reflection *h*.

The structure of *SpLdcB* was solved by SAD using the Se-Met dataset in SHELX (Sheldrick, 2007) as implemented in the HKL2MAP interface (Pape and Schneider, 2004). Analysis of the data using the SHELXC module of SHELX indicated that the anomalous signal extended to approximately 1.9 Å ($d''/\text{sig} > 0.80$) (**Figure 3.7**).

Resl.	Inf	- 8.0	- 6.0	- 5.0	- 4.0	- 3.5	- 3.0	- 2.5	- 2.3	- 2.1	- 1.9	- 1.74
N(data)	245	329	403	898	891	1571	3058	2050	2883	4221	4129	
Chi-sq	1.12	1.06	1.11	0.94	0.94	1.04	1.03	1.04	1.04	1.06	0.95	
< <i>I</i> / σ <i>I</i> >	136.8	118.0	105.0	117.6	98.7	70.5	36.7	21.9	14.7	7.7	2.5	
%Complete	92.1	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.8	81.9	
< <i>d</i> ''/ σ >	14.27	11.33	8.67	6.99	5.49	3.97	2.48	1.66	1.33	1.02	0.82	
CC(anom)	99.7	99.2	98.0	97.1	94.6	90.2	75.1	57.3	38.3	19.6	-2.3	
For zero signal < <i>d</i> ''/ σ > and < <i>d</i> ''/ σ > should be about 0.80												

Figure 3.7: The output of SHELXC. The anomalous signal extended to around 1.9 Å.

A total of six potential selenium sites were identified by SHELXD, with two that had an occupancy of greater than 0.5 (**Figure 3.8**). A total of three methionines are present within *SpLdcB*, suggesting one molecule was present within the asymmetric unit, consistent Matthews' coefficient calculations (Matthews, 1968). Inspection of the electron density maps produced by SHELXE, displayed in COOT (Emsley *et al.*, 2010), in addition to the 'contrast vs cycle' output of SHELXE, indicated that the original hand of the heavy atom substructure was correct (**Figure 3.9**).

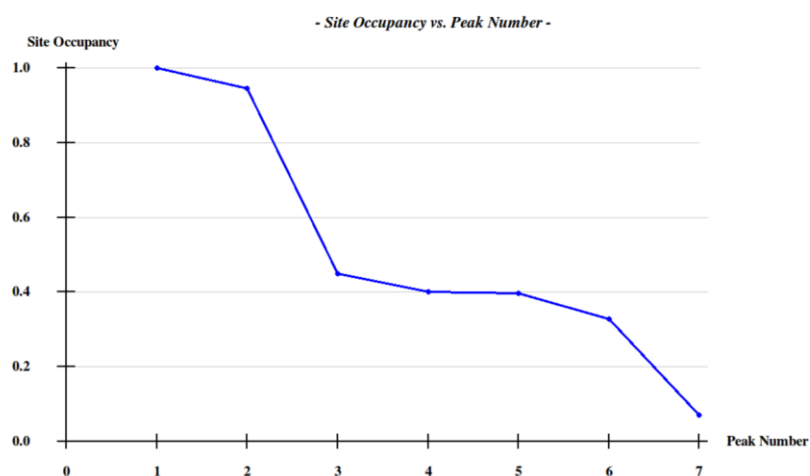


Figure 3.8: The output of SHELXD. The rapid drop-off in site occupancy following the identification of two Se peaks suggested that two selenium atoms were found within the asymmetric unit, corresponding to one protein monomer.

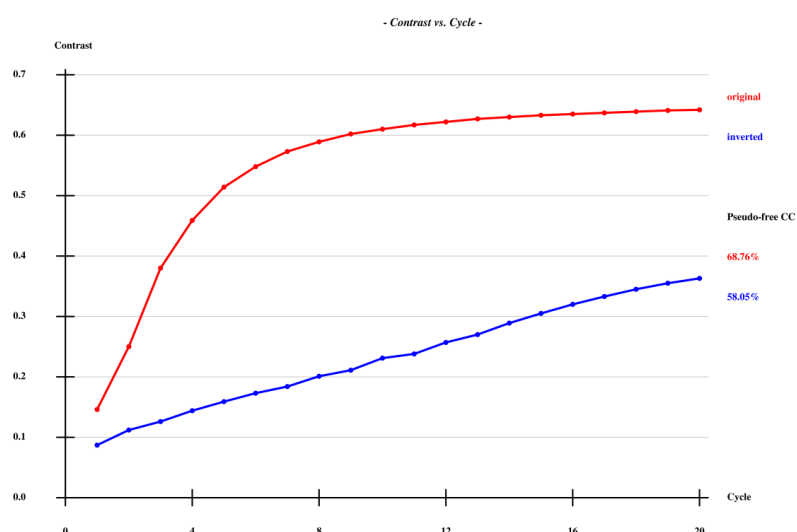


Figure 3.9: The output of SHELXE. Following density modification by SHELXE produces a strong contrast between the solvent and non-solvent areas on the map in the correct hand.

Following identification of the correct hand, initial model building was performed by PHENIX.AUTOSOLVE (Adams *et al.*, 2011), which successfully positioned 180 out of 185 amino acids. The placement of the remaining amino acids, in addition to the zinc ion and solvent atoms, was performed in COOT and the fit of the model to the experimental electron density improved through cycles of manual (re)building in COOT interspaced by restrained refinement in PHENIX.REFINE (Afonine *et al.*, 2012) until convergence. Final refinement statistics can be found in **Table 3.3**.

Table 3.3: Summary of X-ray data refinement statistics for *SpLdcB*.

	<i>SpLdcB</i>
PDBid	4OX5
Resolution (Å)	20.0-1.80
No. Reflections	25656 (1157)
$R_{\text{work}}/R_{\text{free}}^a$	0.191/0.253
No. atoms	
Protein	1476
# proteins/AU	1
Ion/Ligand	3/22
Water	197
B factors (Å ²)	
Protein	41.8
Ion/Ligand	36.9/50.7
Water	48.2
R.m.s deviations	-
Bond lengths (Å)	0.018
Bond angles (°)	1.96
Ramachandran	
favoured (%)	96.7
allowed (%)	100

* Highest resolution shell shown in parentheses.

^a $R_{\text{work}}/R_{\text{free}} = \sum |F_{\text{p}}^{\text{obs}} - F_{\text{p}}^{\text{calc}}| / \sum F_{\text{p}}^{\text{obs}}$, where $F_{\text{p}}^{\text{obs}}$ and $F_{\text{p}}^{\text{calc}}$ are the observed and calculated structure factor amplitudes, respectively.

The final model of *SpLdcB* consisted of residues 56 - 237, with only residues 135, 136 and 238 of the clone missing in the model, due to disorder. The electron density maps were of sufficient quality to also permit the building of residues -4 to -1 (GSHM), present from the pET28a vector. A Ramachandran plot, on which main chain phi angles are plotted against main chain psi angles, revealed the presence of 0 Ramachandran outliers; 96.7 % lie within the favoured regions, with 3.3 % found in allowed regions (**Figure 3.10**, overleaf).

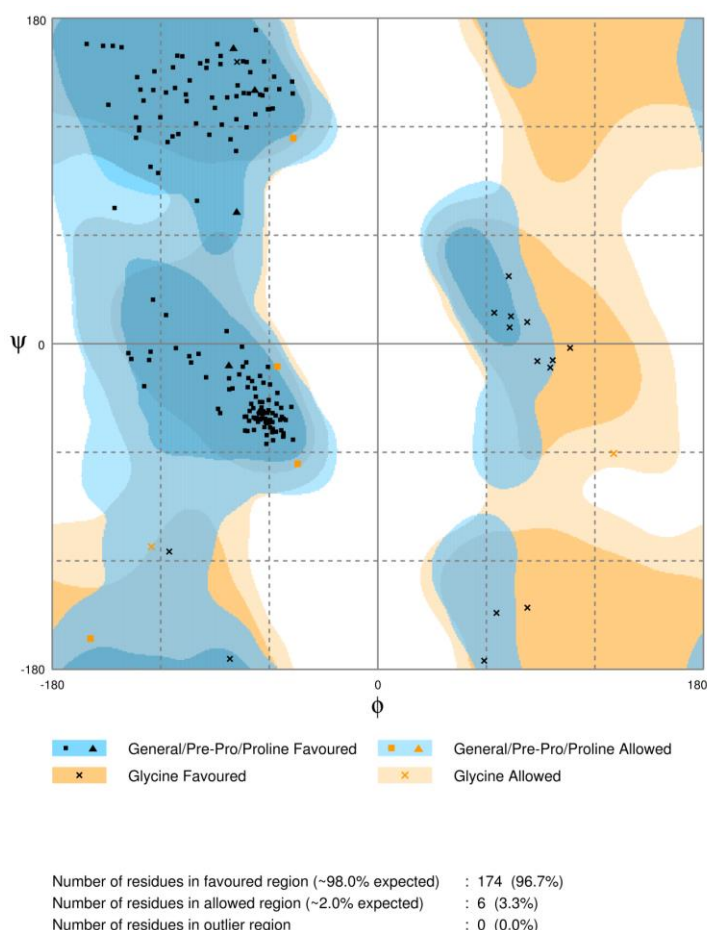


Figure 3.10: Ramachandran plot of the Se-Met *SpLdcB* structure. Black shapes represent residues in favoured regions, orange shapes represent those residues in allowed regions. This plot was generated using the RAMPAGE webserver (Lovell *et al.*, 2003) (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>)

3.3.4 Data processing and molecular replacement of *BsLdcB*

All attempts at improving the quality of *BsLdcB* crystals failed, with the protein consistently crystallising as clusters of needles unsuitable for diffraction studies. However, inspection of the *BsLdcB* crystallisation screens, containing a 4:1 molar excess of zinc, after a period of 6 months revealed the presence of a cluster of plates originating from a single nucleation point (**Figure 3.11**) in the JSCG+ screen condition C6 (0.1 M phosphate/citrate buffer, pH 4.2 and 40% PEG 300). The cluster was harvested in paratone-N oil as described in **Section 3.2.2**.

A native dataset of *BsLdcB* was collected successfully at Diamond Light Source on beamline IO2. A total of 900 images were collected with 0.1° oscillations and 0.1 s exposure per image (**Figure 3.12**, overleaf). Multiple diffraction patterns were found to be present in *BsLdcB* diffraction data (**Figure 3.12**). The dominant lattice was successfully

isolated from the diffraction pattern and data were integrated with XDS (Kabsch, 2010) and reduced with SCALA (Evans, 2005). The crystals were assigned to the spacegroup $P2_12_12_1$ with unit cell dimensions of $a = 42.7$, $b = 53.7$, $c = 102.4$ Å, based upon the systematic absence of reflections where $h+k+l=2n-1$ and statistics following scaling. Data collection statistics are summarised in **Table 3.4**.

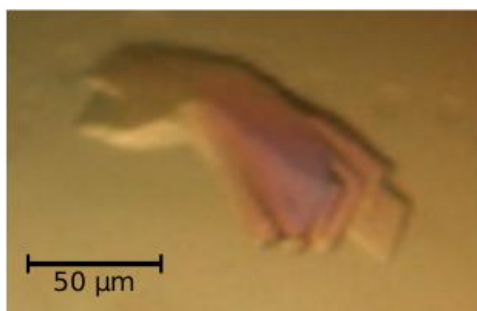


Figure 3.11: *BsLdcB* plate-like crystals.

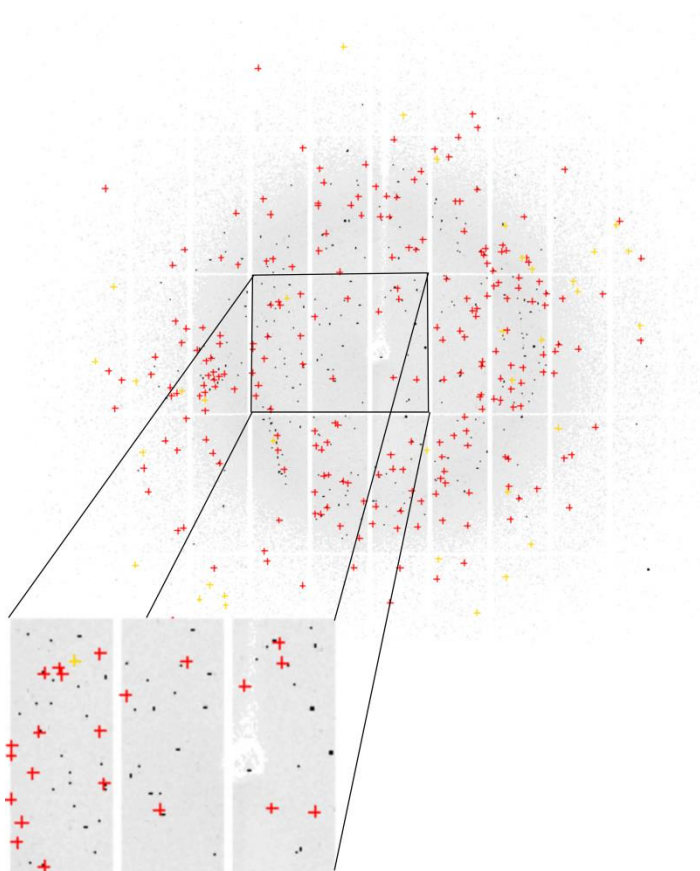


Figure 3.12: An example diffraction pattern produced by *BsLdcB*. The crystal diffracted to 1.7 Å. Multiple lattices are present in the pattern, with spots belonging to the prominent lattice shown as red crosses. Predicted spots are shown as yellow crosses. Prediction and spot-finding was performed by Mosflm (Leslie, 2006), through the graphical interface iMosflm (Battye *et al.*, 2011).

Table 3.4: Summary of X-ray data collection statistics for *BsLdcB*.

	<i>BsLdcB</i>
PDBid	4OX3
Data collection	
Space Group	$P2_12_12_1$
Unit cell dimensions	
a, b, c (Å)	42.7, 53.7, 102.4
α, β, γ (°)	90, 90, 90
Wavelength (Å)	0.97
Resolution (Å)	47.6-1.70
	(1.73-1.70)*
Multiplicity	3.2 (1.7)
R_{merge}^a	0.072 (0.235)
$I/\sigma I$	11.9 (2.4)
Completeness (%)	95.7 (89.3)

^a $R_{\text{merge}} = \sum_h \sum_i |I_i(h) - \langle I(h) \rangle| / \sum_h \sum_i I_i(h)$, where $I_i(h)$ and $\langle I(h) \rangle$ are the i th and mean measurement of the intensity of reflection h .

Matthews' coefficient calculations predicted that 1 molecule of *BsLdcB* was present per asymmetric unit, with a solvent content of 47.1%. The structure of *SpLdcB* was used as a molecular replacement model to solve the structure of *BsLdcB* in PHASER (McCoy et al., 2007), with 1 molecule found per asymmetric unit.

The correctly-positioned protein model was subjected to one round of rigid body refinement in Refmac5 (Winn et al., 2011) before cycles of manual building in COOT and restrained refinement in PHENIX.REFINE (Afonine et al., 2012). The final model of *BsLdcB* consisted of all cloned residues (61 – 273) in addition to a methionine (-1) present from the pET28a vector. A Ramachandran plot reveals the presence of 0 Ramachandran outliers; 98.1 % lie within the favoured regions, with 1.9 % found in allowed regions (**Figure 3.13**). Final refinement statistics can be found in **Table 3.5**.

Table 3.5: Summary of X-ray data refinement statistics for BsLdcB.

	BsLdcB
PDBid	4OX3
Resolution (Å)	47.6-2.0
No. Reflections	35526 (1057)
$R_{\text{work}}/R_{\text{free}}^a$	0.176/0.247
No. atoms	
Protein	1753
# proteins/AU	1
Ion/Ligand	1/5
Water	230
B factors (Å ²)	
Protein	16.0
Ion/Ligand	10.7/23.8
Water	31.6
R.m.s deviations	
Bond lengths (Å)	0.015
Bond angles (°)	1.76
Ramachandran	
favoured (%)	98.1
allowed (%)	100

* Highest resolution shell shown in parentheses.

^a $R_{\text{work}}/R_{\text{free}} = \sum |F_{\text{p}}^{\text{obs}} - F_{\text{p}}^{\text{calc}}| / \sum F_{\text{p}}^{\text{obs}}$, where $F_{\text{p}}^{\text{obs}}$ and $F_{\text{p}}^{\text{calc}}$ are the observed and calculated structure factor amplitudes, respectively.

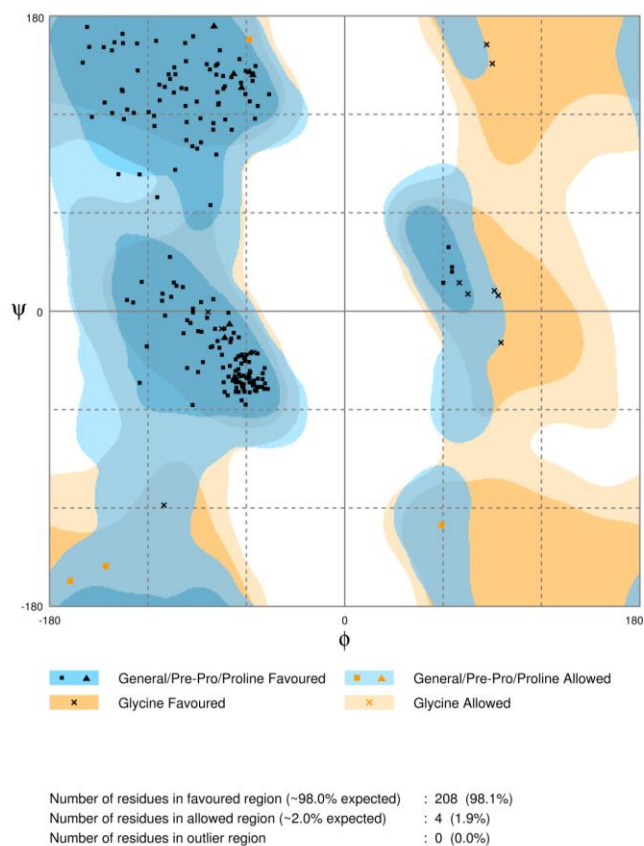


Figure 3.13: Ramachandran plot of the BsLdcB structure. Black shapes represent residues in favoured regions, orange shapes represent those residues in allowed regions. This plot was generated using the RAMPAGE webserver (Lovell *et al.*, 2003) (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>)

3.4 The structure of LdcB

3.4.1 The structures of *Sp*LdcB and *Bs*LdcB

The LdcB enzymes share a very similar overall fold with a core RMSD of 1.8 Å over 154 Cα atoms; both enzymes are single domain proteins comprising two sub-domains separated by a V-shaped cleft (**Figure 3.14 A+B**). The first subdomain comprises a four-stranded, anti-parallel β-sheet (β3-6) that opposes the cleft, and the outside face of the β-sheet is flanked by three α-helices (α1, α4, α6). The second subdomain comprises an N-terminal, two-stranded, anti-parallel β-sheet (β1-2), somewhat discrete from the four helical bundle (α2, α3, α5, α7) that makes up the rest of the subdomain. Neither subdomain is formed by a contiguous, linear part of the primary sequence; the two subdomains are connected by six loops (β3-α1, β4-α2, α3-β5, β5-α5, α5-β6, α6-α7).

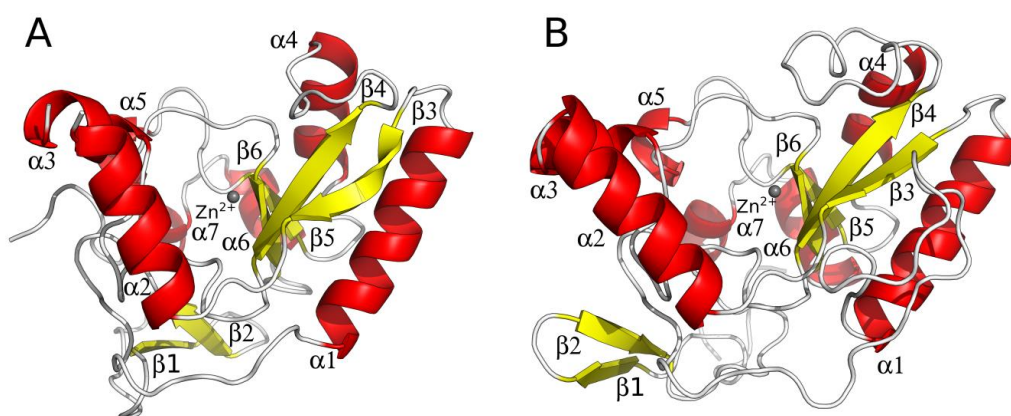


Figure 3.14: Structure of LdcB. Cartoon representation of *Sp*LdcB (**A**) and *Bs*LdcB (**B**); α-helices (red) and β-strands (yellow) are numbered. The loops between secondary structure elements are coloured white. The bound zinc ion is shown as a grey sphere. Unless otherwise stated, this, and all subsequent figures, were produced with PyMOL (Schrodinger, 2010)

The cleft between subdomains is 15-18 Å wide at the top, 15 Å deep and narrows down to the base where a Zn²⁺ ion is located, which is coordinated in a tetrahedral manner by three amino acids (**Figure 3.15**, overleaf). In *Sp*LdcB, the liganding residues are His153, Asp160 and His207 (**Figure 3.15**, overleaf); the equivalent residues in *Bs*LdcB are His185, Asp192 and His244 (**Figure 3.15**, overleaf). These zinc-ligands are universally conserved in LdcB sequences (**Figures 3.16 and 3.17** overleaf), consistent with an important functional role in these enzymes. The zinc-ligands are themselves hydrogen bonded by a set of conserved amino acids. In *Sp*LdcB (corresponding *Bs*LdcB residues are labelled

in parentheses) Gly118 (Gly150) interacts with His153 (His185) while Glu152 (Glu184) hydrogen bonds with His207 (His244), providing structural rigidity to the active site (**Figure 3.15**). In addition to these Zn^{2+} binding amino acids, the active site contains Glu204 (Glu241) which might be capable of acting as an acid/base during catalysis, and Arg120 (Arg152), which, with its proximity to the zinc, could stabilise a negatively-charged intermediate reaction state (**Figure 3.15**). These active site residues are strictly conserved in various LdcB and VanY family enzymes (**Figure 3.17**), with the exception of Gly118 equivalents, which are found to be replaced by threonine in two sequences. This mutation is not surprising as it is the carbonyl oxygen of the glycine that hydrogen bonds to the histidine, and thus the hydrogen bonding network is conserved despite this substitution. Possible catalytic mechanisms involving these residues are suggested and expanded upon in **Section 4.5**.

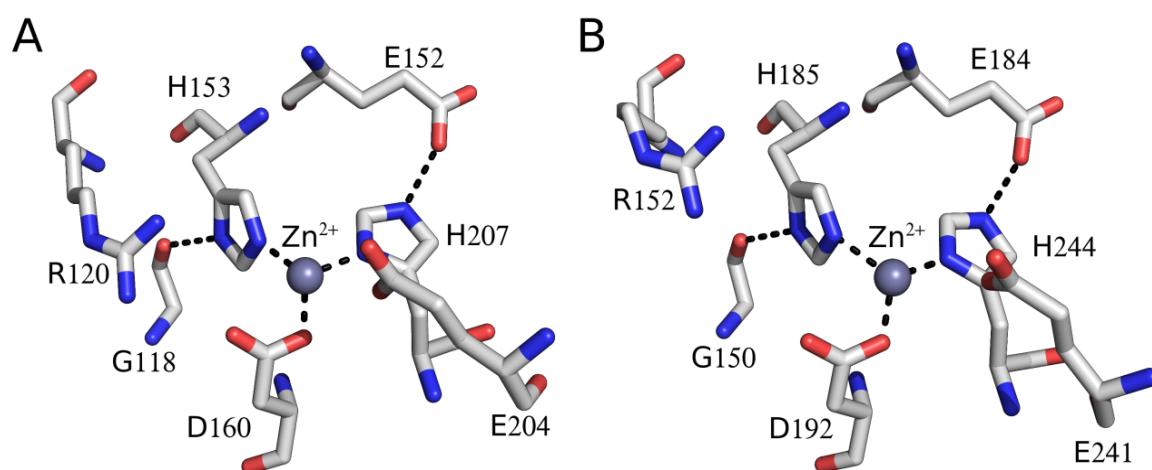


Figure 3.15: Structure of the LdcB active site. Stick representations of *Sp*LdcB (**A**) and *Bs*LdcB (**B**) active sites. A zinc ion is coordinated by two invariant histidines and a well-conserved aspartate. The two histidines are hydrogen bonded to either a glutamate or glycine, providing rigidity to the active site. Carbons are shown in white and the bound zinc ion is shown as a grey sphere.

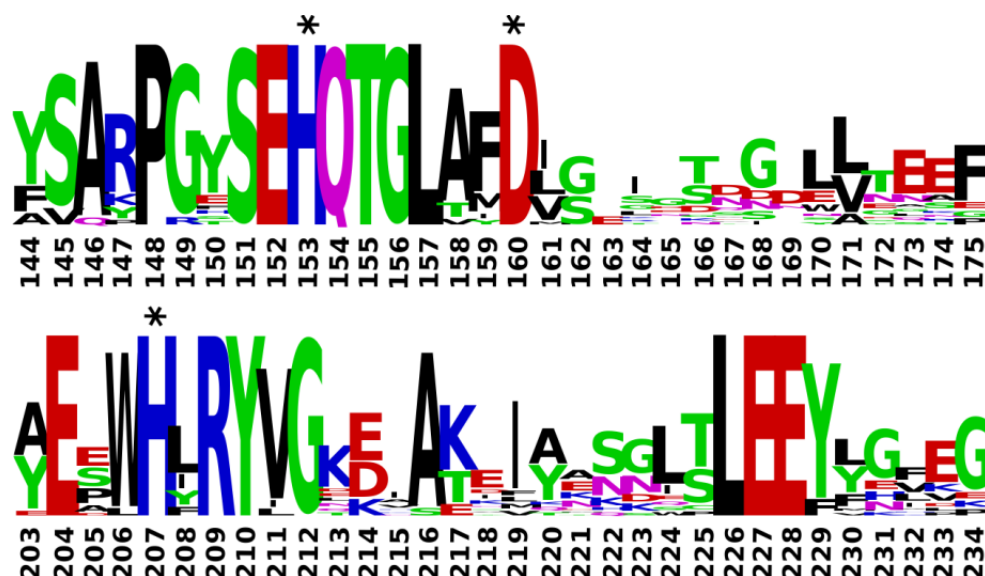


Figure 3.16: Conservation of zinc-coordinating residues. Weblogo (Crooks *et al.*, 2004) motif showing the conservation of the zinc-coordinating residues (indicated by an asterisk) in LdcB sequences from a range of organisms (*S. pneumoniae*, *S. pseudopneumoniae*, *S. mitis*, *S. infantis*, *S. oralis*, *G. sanguinis*, *L. fusiformis*, *G. adiacens*, *G. elegans*, *Fermicutes bacterium M10-2*, *K. huakuii*, *B. subtilis*, *B. anthracis*, *L. lactis*, *C. difficile*, *S. aureus* and *L. monocytogenes*). Amino acids are labelled based upon the sequence of *SpLdcB*. Polar amino acids (G, S, T, Y, C, Q, N) are shown in green, basic residues (K, R, H) in blue, acidic residues (D, E) in red and hydrophobic (A, V, L, I, P, W, F, M) amino acids are black. Sequences were aligned in MUSCLE (Edgar, 2004).

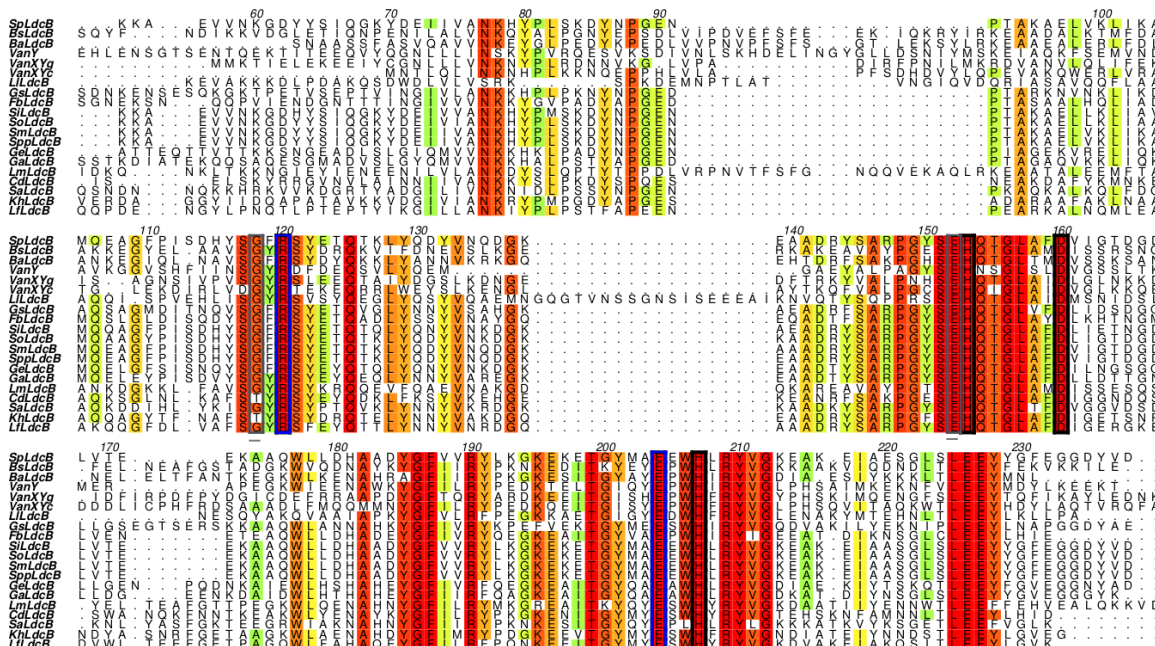


Figure 3.17: Sequence alignment of a range of LdcB sequences with VanY, VanXYg and VanXYc. Sequence conservation is shown as a rainbow (cyan>red), with red indicating highest conservation. All metal binding residues are shown in black boxes, their hydrogen bonding partners are shown in grey boxes with grey underline and residues hypothesized to be involved in catalysis are shown in blue boxes. Residues are numbered in reference to the *SpLdcB* sequence. *Spp* = *S. pseudopneumoniae*, *Sm* = *S. mitis*, *Si* = *S. infantis*, *So* = *S. oralis*, *Gs* = *G. sanguinis*, *Lf* = *L. fusiformis*, *Ga* = *G. adiacens*, *Ge* = *G. elegans*, *Fb* = *Fermicute bacterium M10-2*, *Kh* = *K. huakuii*, *Li* = *L. lactis*, *Cd* = *C. difficile*, *Sa* = *S. aureus* and *Lm* = *L. monocytogenes*).

3.4.2 Identification of native-bound ligands

The active site of *Sp*LdcB contains unattributed residual electron density that occupies the position of the fourth ligand in the tetrahedral coordination sphere of the zinc (Figure 3.18). To use the nomenclature common to proteases, the unattributed electron density extends as far as the S1' subsite, the location ordinarily occupied by the scissile amino acid. The ligand is stabilised in this position by van der Waals' contacts between the alanyl methyl group and the phenolic side chains of Tyr191 and Tyr201, and by hydrogen bonds to the side chains of Glu204 and the main chain amide nitrogen of Ala146 (Figure 3.18). Despite the high resolution of the *Sp*LdcB diffraction data, it was not possible to determine the identity of the ligand from the shape of the electron density alone, which is of the approximate size expected of a dipeptide. Although the crystallisation condition did contain a number of amino acids, it did not contain any dipeptides nor any other molecules of the mass to account for the electron density, therefore the ligand is likely to have co-purified with the enzyme from *E. coli* during its overexpression.

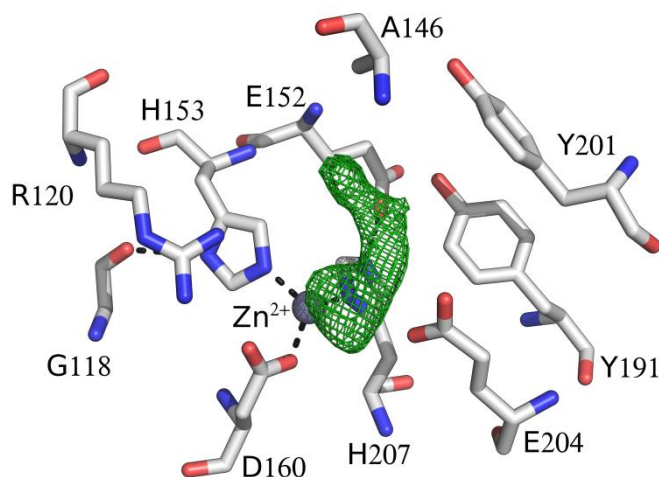


Figure 3.18: Unknown density in the LdcB active site. Electron density corresponding to an unknown ligand (green mesh) was observed in the active site of *Sp*LdcB. Y191 and Y201 appear to stabilise the ligand, and E204 and A146 are within hydrogen bonding distance. *Sp*LdcB is shown in stick representation with white carbons, the bound zinc ion is shown as a grey sphere. Electron density maps are 2Fo-Fc and are contoured to 3.0 σ .

Two complementary methods were employed in an attempt to identify the ligand. First, purified *SpLdcB* was submitted for native and denatured mass spectrometry. By comparing the apparent mass of the native protein, which includes the masses of any bound ligands, with that of the denatured protein, the mass of any bound ligands could be deduced by simple subtraction. Databases such as the NIST chemical database (<http://webbook.nist.gov>) were queried with the mass differences obtained in a search for likely chemical identities.

SpLdcB was dialysed into 10 mM ammonium acetate buffer pH 8 prior to being concentrated to 10 mg/ml and submitted to for native and denatured electrospray mass spectrometry at the Mass Spectrometry facility, Astbury Centre for Structural Molecular Biology at the University of Leeds, UK, by Dr James Ault. *SpLdcB* was analysed on a Platform II mass spectrometer (Waters). The denatured protein sample produced a mass of 21170.5 Da (**Figure 3.19**, overleaf), consistent with that of *SpLdcB* (21170.2 Da, calculated by the ProtParam webserver (Gasteiger *et al.*, 2005)). The native protein produced a spectrum with a large peak at 21236.3 Da, corresponding to *SpLdcB* with bound zinc ion of 65.9 Da ($21236.3 - 21170.2 = 66.1$ Da). Additional peaks with masses of 21245.4 and 21335.1 Da in the native *SpLdcB* spectrum could correspond to molecules that fulfil unmodelled electron density.

The unknown ligand is observed to coordinate to *SpLdcB* via the active site zinc (**Figure 3.18**) and thus the mass of *SpLdcB*:zinc (21236.3 Da) was subtracted from the remaining mass peaks (21245.4 and 21335.1 Da) to determine the mass of the ligand. This resulted in two mass differences; 9 Da, far too small to account for the observed density, and 99 Da. 99 Da does not correspond to any dipeptides (mass ranges of 115 Da [Gly-Gly] – 373 Da [Trp-Trp]), or to any amino acids. A search of NIST with a mass difference of 99 Da produces a list of 195 candidates, with masses that range from 98.9 – 99.16 Da. However, none of the proposed chemicals provide sufficient explanation for the observed electron density. Expanding the search further, to include molecules with masses ranging from 98 – 101 Da, produces a list of over 400 candidates and it quickly becomes unfeasible to interrogate the list effectively.

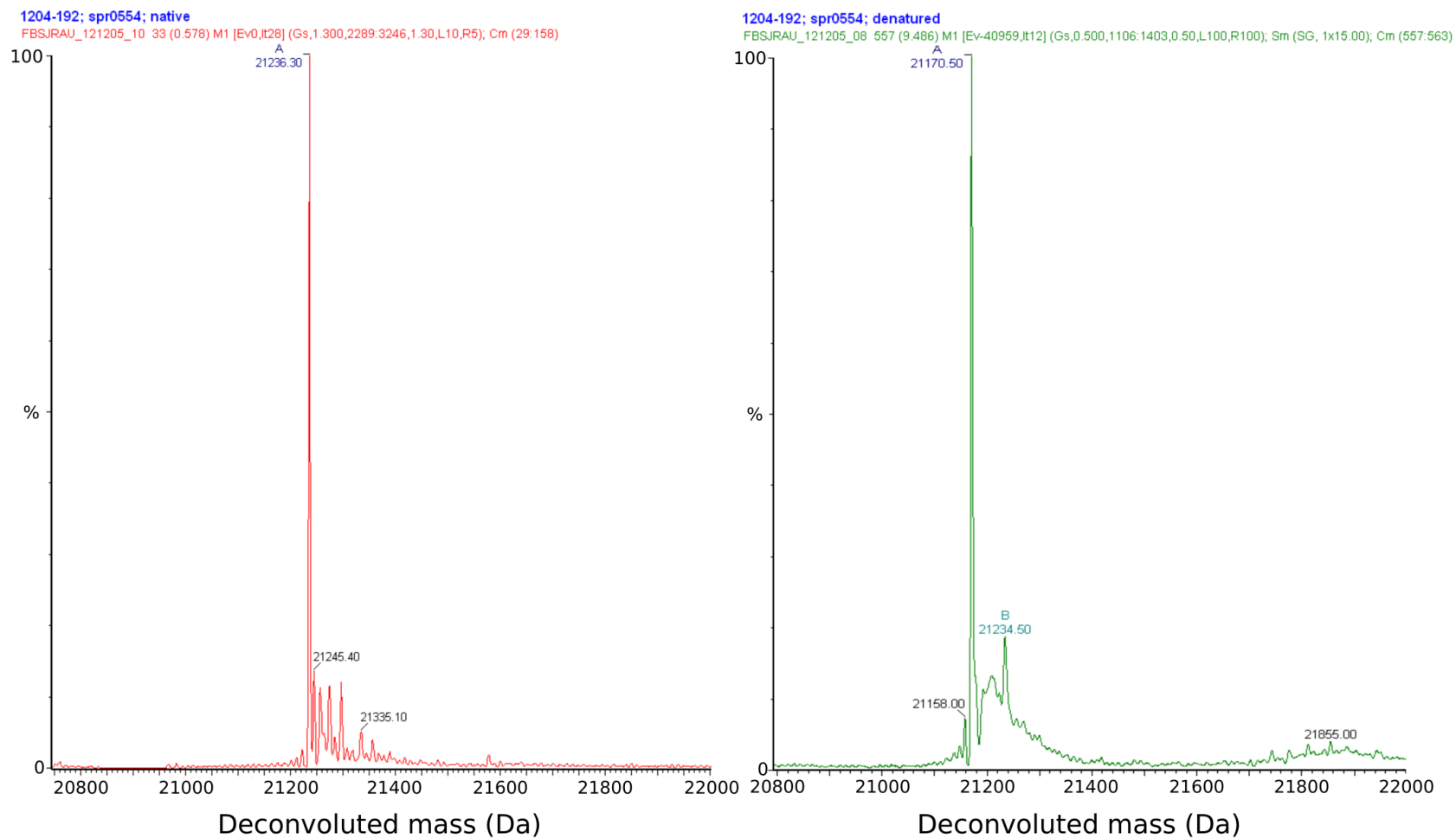


Figure 3.19: Electrospray mass spectra of *SpLdcB* native (top) and denatured (bottom). The differences seen between the mass of the denatured and native *SpLdcB* do not correspond to an amino acid or dipeptide. Electrospray mass spectroscopy was performed by Dr James Ault, Leeds University.

The second method employed to identify the unknown ligand assumed that the ligand was an amino acid or dipeptide. This assumption is not unwarranted as the reaction products of LdcB are an amino acid and a tripeptide; and a dipeptide could easily bind to an active site designed to accommodate tetrapeptides. The ligand, once theoretically extracted from *SpLdcB* by thermal denaturation and centrifugal filtration, was chemically linked to FDAA (1-fluoro-2,4-dinitrophenyl-5-L-alanine amide) (**Section 2.5.2.2**). The resultant diastereomer was then loaded onto a C-18 reverse phase HPLC column and eluted by an acetonitrile gradient. Two peaks, in addition to two, large, free FDAA peaks, were observed. The retention time of the two peaks was compared to that of control FDAA-linked amino acids and dipeptides, with one peak matching that of FDAA-L-alanine and the other FDAA-D-alanine (**Figure 3.20**).

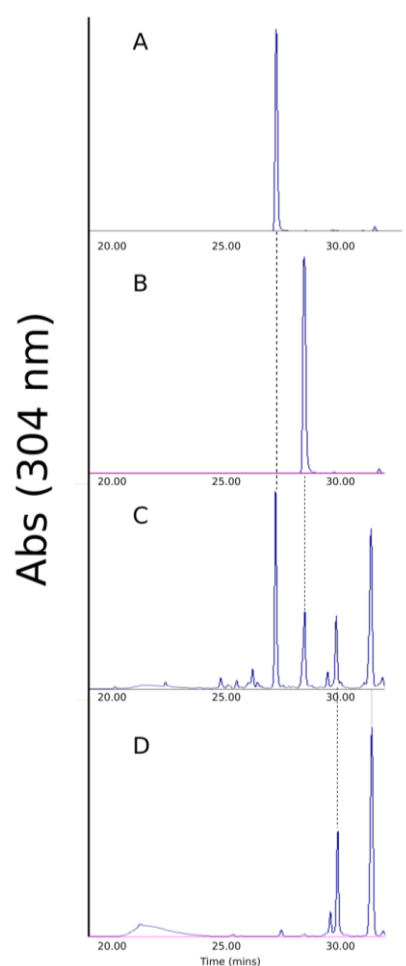


Figure 3.20: Reverse-phase HPLC chromatograms for FDAA-linked L-alanine (A), D-alanine (B) and unknown ligand (C) and FDAA by itself (D). All significant peaks in panel C correspond to peaks in the other three panels; either FDAA-linked L-alanine (1st peak), FDAA-linked D-alanine (2nd peak) or unlinked FDAA (peaks 3 and 4).

Although alanine cannot adequately explain the entirety of the electron density observed at the S1 and S1' subsites, both stereoisomers fit well into the portion of the electron density at S1', where the terminal D-alanine of the cell wall tetrapeptide would be located during catalysis (**Figure 3.21**), prior to and following refinement. Ultimately, the D-isomer of alanine was modelled as it is the enzymatic product of these enzymes. Unfortunately, the residual electron density coordinating directly to the zinc ion (purple) could not be modelled satisfactorily with any ligand arising from the enzymatic activity, the crystallisation milieu or purification media, and consequently the remaining electron density has been left unmodelled (**Figure 3.21**).

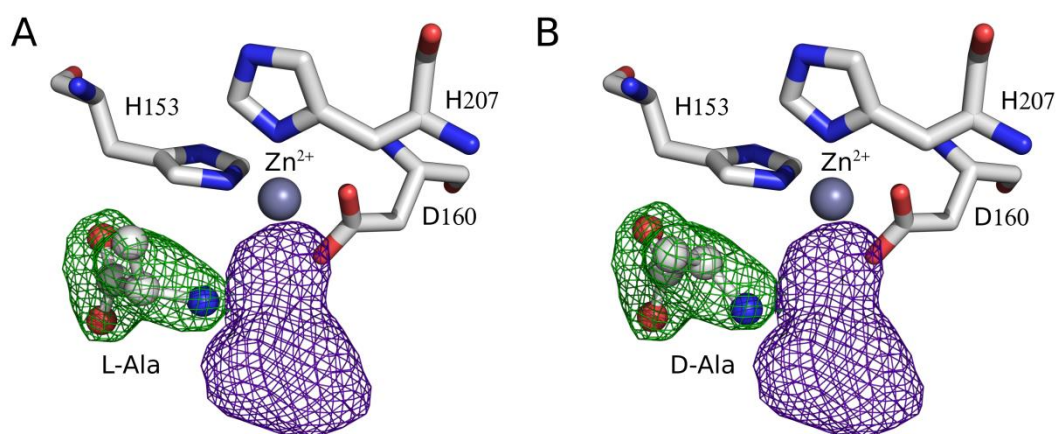


Figure 3.21: D- and L-Alanine fitting into the unknown *SpLdcB* density. Both D- and L- alanine fit equally well in the unknown density (green mesh). The identity of the second electron density mesh (purple) could not be determined and has been left unmodelled. *SpLdcB* is shown in stick representation with white carbons, the bound zinc ion is shown as a grey sphere. Electron density maps are 2Fo-Fc and are contoured to 3.0 σ .

As mentioned previously, the shape of the electron density (both green and purple meshes) appears to be of a dipeptide. In addition, the proximity of the amino group of the D-alanine, within 1.55 Å of the unmodelled density, indicates that the D-Alanine may be covalently linked to another molecule (amino acid or otherwise). In addition, the size of the unmodelled electron density is similar to that of the density that corresponds to the D-alanine, albeit a different shape, suggesting that the unidentified molecule is likely around 90-110 Da in mass.

Residual electron density was also identified in the active site of *BsLdcB*, in the vicinity of the catalytic zinc ion. This density was fulfilled by a phosphate anion, a component of the crystallisation condition, which refined satisfactorily (**Figure 3.22**).

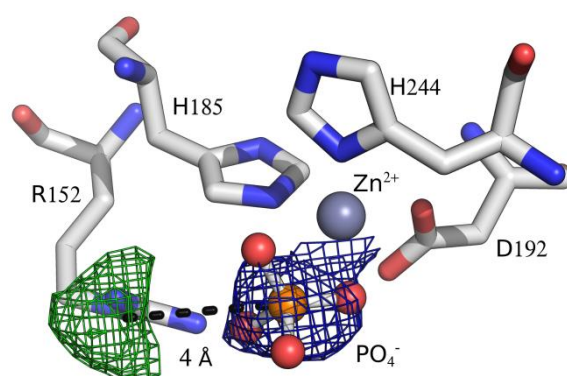


Figure 3.22: Density of a phosphate ion in the *BsLdcB* active site. Density that corresponds to a phosphate ion was found in the active site of *BsLdcB* (blue mesh). Additional density was located within 4 Å of the first phosphate (green mesh), with electrostatic repulsion making it unlikely that the unknown density corresponded to a second phosphate. *BsLdcB* is shown in stick representation with white carbons, the bound zinc ion is shown as a grey sphere and the phosphate ion in ball-and-stick representation. Electron density maps are 2Fo-Fc and are contoured to 3.0 σ

There is a further unmodelled electron density feature, this time in the S1' subsite, the approximate size and shape of a phosphate anion. However, the proximity of this second density feature, 4 Å from the first, makes it unlikely that both are phosphate anions. Whilst it is possible that the presence of the nearby Arg152 could stabilise the electrostatic repulsion that would be experienced by two phosphates, we have chosen to leave the second electron density region unmodelled in the final structure.

3.4.3 Comparing the LdcB structures

During the course of this project, LdcB from *Bacillus anthracis* (*BaLdcB*) was crystallised and its structure solved by the Center for Structural Genomics of Infectious Disease (CSGID) (Hoyland *et al.*, 2014). Like *SpLdcB* and *BsLdcB*, *BaLdcB* is comprised of a single domain, with two sub-domains separated by a V-shaped cleft, though the second sub-domain is less well-structured than in *BsLdcB* and *SpLdcB* and as a consequence the β 1- β 2 sheet is not formed (**Figure 3.23**).

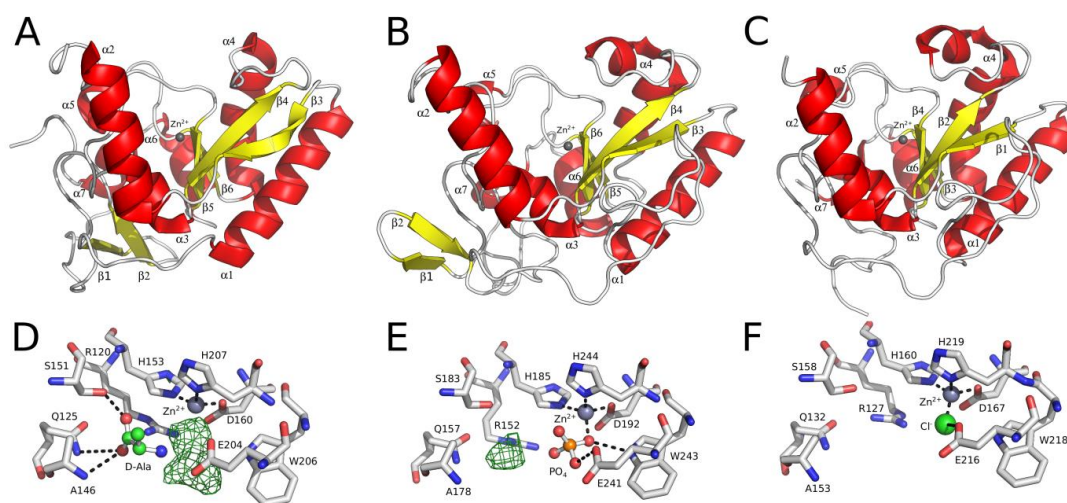


Figure 3.23: Comparing the three LdcB structures. (A-C) Cartoon representations of *SpLdcB* (A), *BsLdcB* (B), and *BaLdcB* (C); α -helices (red) and β -strands (yellow) are numbered. The loops between secondary structure elements are coloured silver. The bound zinc ion is shown as a grey sphere.

(D-F) The active sites of *SpLdcB* (D), *BsLdcB* (E), and *BaLdcB* (F). A zinc ion is coordinated in a tetrahedral manner in the structures by two invariant histidines and a well-conserved aspartate. The residual difference density ($2F_o - F_c$) features are shown contoured at a level of 3σ and hydrogen bonds are drawn as dashed black lines.

SpLdcB, *BsLdcB* and *BaLdcB* share pairwise sequence identities of $\sim 30\%$, and they superimpose on ~ 155 common C α atoms with an RMSD of $1.5\text{ \AA} - 1.8\text{ \AA}$. The major differences between the LdcB structures are limited to the rather flexible and unconstrained N- and C-termini and the occasional longer loop in one structure in comparison to the others, due to sequence insertions. The protein atoms of the active sites of all three LdcB enzymes are, to all extents and purposes, indistinguishable (**Figure 3.23, D-F**). A further *BaLdcB* structure has been solved that contains no bound zinc ion, but there are negligible differences between it and its counterpart with zinc bound (Hoyland *et al.*, 2014). The LdcB enzymes therefore do not appear to require zinc to correctly fold, though the presence of the metal ion is required for catalysis (**Section 4.3**) (Hoyland *et al.*, 2014).

Interestingly, the structure of *SpLdcB* helps to explain the effect zinc has upon crystallisation of these enzymes. When zinc binding of LdcB was assessed in solution, the excess zinc was removed by size exclusion. By contrast, *SpLdcB* was crystallised with a 4:1 excess of zinc and so each molecule of *SpLdcB* binds one zinc in the active site and shares two further zinc ions with neighbouring LdcB proteins in the crystal lattice (**Figure 3.24, overleaf**). The two surface zinc-binding sites differ in that the first site (pink:green contact in **Figure 3.23**) is comprised of 3 amino acid ligands; His-2 and Glu56

from the pink monomer and His79 from the green molecule, and a water is present as the final ligand. The second site, between the green and blue molecules (**Figure 3.24**, overleaf), is comprised of His181 and Asp184 from the green monomer, a serendipitous tris ligand from the crystallisation milieu and His115 from the blue molecule. In a further act of serendipity, histidine -2 is from the pET28a vector used for overexpression; not only does the crystal contact involving His-2 improve the quality of the crystals, it stabilises the tag and allows the selenomethionine at position -1 to contribute to the anomalous signal, accounting for the 3rd site with 50% occupancy identified by SHELXD during the selenium heavy atom substructure search (**Figure 3.8**). Corroborating this, the B-factor of Mse-1 is almost double that of the other Se-Met residues of *SpLdcB* (Mse-1 = 59.0 Å², Mse105 = 24.2 Å², Mse202 = 33.9 Å²).

The intermolecular zinc binding sites, although not possessing any physiological relevance, are analogous to common zinc binding sites. The site between the pink and green monomers is comprised of a two histidines and a glutamate in the same arrangement as the metal binding site found in the canonical carboxypeptidase, pancreatic carboxypeptidase A (Christianson and Lipscomb, 1989), whilst the site situated between the green and blue monomers is identical to the catalytic zinc site of LdcB, being comprised of two histidines and an aspartate. A tris molecule completes the tetrahedral geometry of the green:blue metal binding site. The *BsLdcB* protein binds only 1 zinc ion per monomer, in the active site, and so the addition of excess zinc failed to promote crystal contacts and enhance crystallisation.

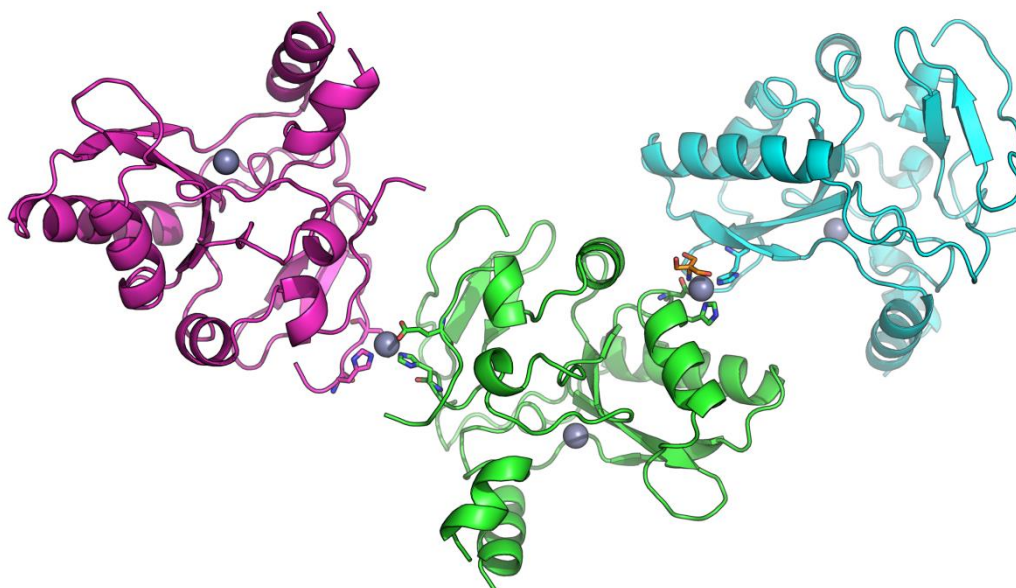


Figure 3.24: Zinc forms crystal contacts between *SpLdcB* monomers. Three neighbouring *SpLdcB* monomers are shown in cartoon ribbon representation. The zinc ions bound by these proteins are shown as grey spheres, with the liganding residues shown as sticks. The tris molecule is shown in stick representation with orange carbons.

3.4.4 Comparison of *SpLdcB* and *BsLdcB* with LAS (lysostaphin, D-Ala-D-Ala metallopeptidases, sonic hedgehog) family proteins

Using *SpLdcB* as a search model in DALI (Holm and Rosenström, 2010), over 50 similar structures were identified in the PDB (**Table 3.6**), including the structure of VanXYc, a dipeptidase/DD-carboxypeptidase; VanXYg, a second dipeptidase/DD-carboxypeptidase; an endolysin from the *Listeria* bacteriophage A500; VanX, a DD-dipeptidase and the N-terminal domain of sonic hedgehog, a protease involved in embryogenesis. Equivalent results were obtained using *BsLdcB* or *BaLdcB* as queries in DALI. The matched structures are all members of the wider LAS (lysostaphin, D-Ala-D-Ala metallopeptidases, sonic hedgehog, (Bochtler *et al.*, 2004)), family of proteases which, although share very little sequence identity, possess markedly similar active sites and core fold (**Figure 3.25**).

Table 3.6: Table summary of a section of DALI results. The structure of *SpLdcB* was used as a search model.

PDBid	Name	Z-score	% identity	Reference
4MUT	VanXYc	18.6	32	Meziane-Cherif <i>et al.</i> , 2014
4F78	VanXYg	17.0	37	Meziane-Cherif <i>et al.</i> , 2014
2VO9	An endolysin from the <i>Listeria</i> bacteriophage A500	9.5	19	Korndörfer <i>et al.</i> , 2008
1R44	VanX	6.6	20	Bussiere <i>et al.</i> , 1998
2WG4	N-terminal sonic hedgehog	5.3	19	Bishop <i>et al.</i> , 2009
1LBU	D-Ala-D-Ala carboxypeptidase from <i>Streptomyces albus</i>	4.4	19	Not published

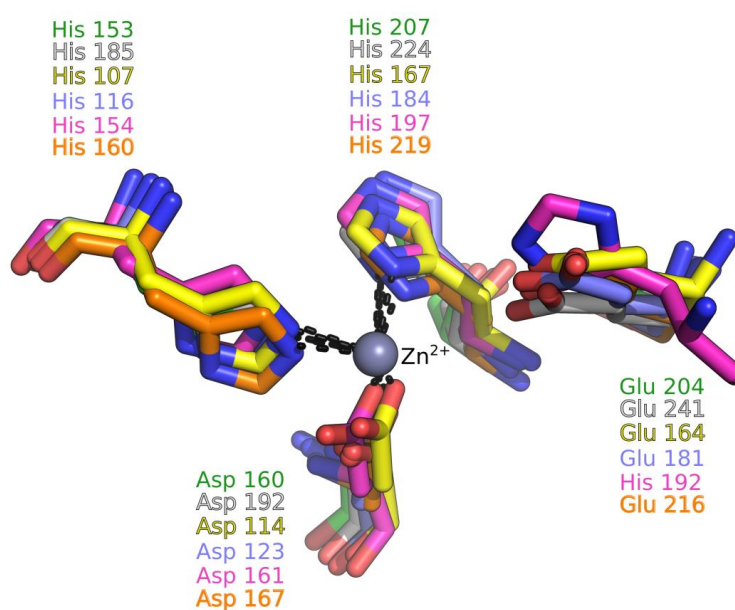


Figure 3.25: A comparison of the active sites of LdcB and other LAS family members. Both *SpLdcB* and *BsLdcB* share an active site closely superimposable on those of other LAS family enzymes. The zinc is co-ordinated by a H-X₍₃₋₆₎-D-X_(not conserved)-H motif. A glutamate that plays an important role in catalysis is also shown; note that in D-Ala-D-Ala carboxypeptidase from *Streptomyces albus*, the equivalent amino acid is histidine, the first amino acid in the H-X-H motif common to many LAS carboxypeptidases. The structures shown, distinguished by the colouring of carbon atoms, are *SpLdcB* (white), *BsLdcB* (green), *BaLdcB* (orange), *Enterococcus faecalis* D,D-dipeptidase/D,D-carboxypeptidase VanXY_G (yellow; PDBid: 4F78), *Enterococcus faecium* aminodipeptidase VanX (blue; PDBid: 1R44) and *S. albus* D-Ala-D-Ala carboxypeptidase (pink; PDBid: 1LBU).

Although the overall folds of the various LAS super-family proteins differ, every member contains a core β -sheet comprised of a minimum of 4 antiparallel β -strands, which acts as a scaffold for the placement of residues used to bind a catalytic zinc ion (Bochtler *et al.*, 2004). These Zn^{2+} ligands are invariably found to be His-Asp-His and can be identified by a H-X(3/6)-D-X(not conserved)-H consensus sequence. The first histidine of the zinc binding motif consistently extends from a loop that connects strands 1 and 2 of the β -sheet ($\beta 3$ and $\beta 4$ in *SpLdcB* and *BsLdcB*, $\beta 1$ and $\beta 2$ of *BaLdcB*) and coordinates the Zn^{2+} ion through the N(ϵ) of its imidazole ring (Bochtler *et al.*, 2004). The orientation of the histidine positions the N(δ) of the imidazole ring within hydrogen bonding distance of a main-chain carbonyl oxygen atom (Gly118 in *SpLdcB*) and its protonated state allows it to act as the hydrogen bond donor. The second residue of the triad, aspartic acid, extends from the second strand of the β -sheet and is either 3 or, in the case of *SpLdcB* and the Van-family enzymes, 6 amino acids downstream of the first histidine (Bochtler *et al.*, 2004). Finally, the third histidine is found at various distances downstream of the first two amino acids, but invariably extends from the 4th strand of the β -sheet (Bochtler *et al.*, 2004), which is $\beta 6$ in *SpLdcB* and *BsLdcB*, $\beta 4$ in *BaLdcB*. Like the first histidine, both nitrogens of the imidazole ring are used to lock the residue in the correct orientation, though the roles of the N(ϵ) and N(δ) are reversed; the N(δ) is used to coordinate the Zn^{2+} ion, resulting in a protonated N(ϵ) that then acts as hydrogen bond donor to a glutamate residue (Glu152 in *SpLdcB*, Glu184 in *BsLdcB*).

In addition to the three protein ligands, the zinc ion is invariably coordinated by a fourth ligand, which tends to be either a water that is likely to act as the nucleophile in the reaction, or a solute from the crystallisation milieu. In the structure of the LAS family LytM proenzyme, an inhibitory loop folds into the active site and results in the fourth coordination ligand being an asparagine from the inhibitory loop (Bussiere *et al.*, 1998; Bochtler *et al.*, 2004; Odintsov *et al.*, 2004). A phosphate anion from the crystallisation media occupies the fourth ligand site in coordinating the Zn^{2+} in *BsLdcB*, whilst an unknown ligand plays this role in *SpLdcB* (**Section 3.4.2**).

3.4.5 Comparing *SpLdcB* and *BsLdcB* with vancomycin resistance enzymes

The LAS family can be divided into two sub-groups based upon a second conserved motif; those that contain a H-x-H motif, in which the first histidine has been shown to be the catalytic histidine, and those that contain the motif E-x-x-H (Bochtler *et al.*, 2004). All three LdcB enzymes possess the E-x-x-H sequence and are therefore placed within the second group, alongside some vancomycin resistance enzymes including the D-Ala-D-Ala dipeptidase VanX, the DD-carboxypeptidase VanY and the two bi-functional enzymes VanXYg and VanXYc, which possess both dipeptidase and carboxypeptidase activities (Meziane-Cherif *et al.*, 2014). Sequence alignments of *SpLdcB* with VanX, VanY, VanXYg and VanXYc using BLAST (Altschul *et al.*, 1990) reveal sequence identities of between 22 – 38 % , with *SpLdcB* sharing highest homology with VanXYg (36%, 79% coverage, E-value 1×10^{-21}), followed by VanXYc (30%, 74% coverage, E-value 1×10^{-21}) and VanY (38%, 47% coverage, E-value 2×10^{-19}). This is not unexpected as all three of these enzymes target very similar substrates, the stem peptides of peptidoglycan. VanX is the least similar, with a sequence identity of only 22%, with 18% coverage (E-value 1.6). Again, this was expected, as VanX is a dipeptidase that cleaves D-Ala-D-Ala dipeptides, rather than a carboxypeptidase.

Superposing the three LdcB homologues with VanXYc and VanXYg in COOT (Emsley *et al.*, 2010) reveals core RMSDs of between 1.5 – 2.0 Å over 146 – 167 aligned C α atoms (**Table 3.6**), with all active site residues aligning almost perfectly. Indeed, LdcB enzymes feature a serine (*SpLdcB*, Ser117; *BsLdcB*, Ser149; *BaLdcB*, Ser124) and glutamine (*SpLdcB*, Gln125; *BsLdcB*, Gln157; *BaLdcB*, Gln132) pair that interact with an arginine (*SpLdcB*, Arg120; *BsLdcB*, Arg152; *BaLdcB*, Arg127) present at the base of the active site to stabilise its conformation, a feature previously thought to be unique to the VanXY enzymes (**Figure 3.26**)(Meziane-Cherif *et al.*, 2014). It is clear that LdcB enzymes have a much more solvent-accessible active site than the VanXY enzymes, a direct consequence of the absence from LdcB of the β 4- α 5/bi-substrate selectivity loop (Meziane-Cherif *et al.*, 2014). Presumably the β 4- α 5/bi-substrate selectivity loop is required by the VanXY enzymes to cap the active site during the hydrolysis of dipeptide substrates, with the loop adopting a more open confirmation to allow the enzymes to be active against pentapeptide substrates (Meziane-Cherif *et al.*, 2014).

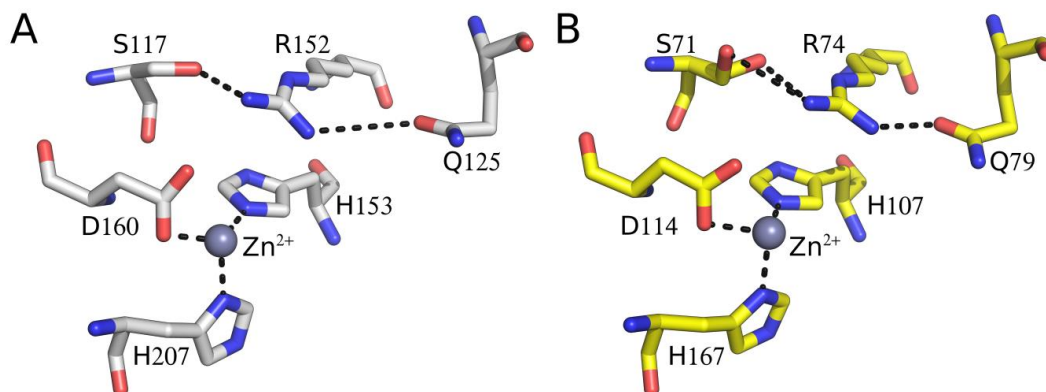


Figure 3.26: A comparison of the active sites of *SpLdcB* and *VanXYg*. The active sites of LdcB and VanXY proteins are almost identical. Both proteins are shown in stick representation, with the carbons of *SpLdcB* shown in white, *VanXYg* is shown in yellow. Both conformations of S71 of *VanXYg* are shown. The zinc is shown as a grey sphere and hydrogen bonds are shown as black dotted lines.

Aside from the bi-substrate selectivity loop, the other notable difference between *VanXYg* and *LdcB* enzymes is the presence of a second domain formed from residues 1–32 and 186–250 in *VanXYg* (**Figure 3.27**). This domain is of unknown function, and is not present in the structure or coding sequence for the other *VanXY* enzyme, *VanXYc*. Sequence homology suggests that this second domain may be present in *VanY*, where it has been hypothesized to be involved in membrane association (Meziane-Cherif *et al.*, 2014). Alternatively, this domain may function in a similar manner to the ‘pedestal’ domain of the class B PBPs such as PBP2b, which is hypothesised to project the catalytic domain of the protein away from the membrane to position it next to its substrate in the cell wall (Macheboeuf *et al.*, 2006), though it should be noted that this second domain shares no sequence homology with the ‘pedestal’ of PBP2b. In either scenario, the function of this domain is clearly lost in *VanXYg*.

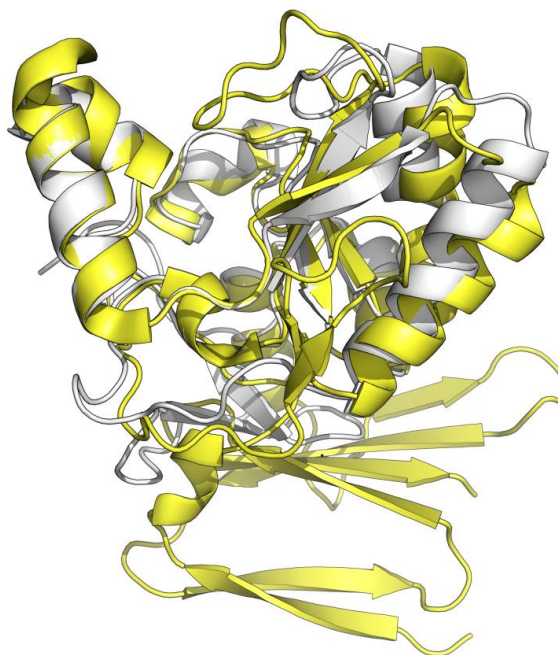


Figure 3.27: SpLdcB superposed onto VanXYg. *SpLdcB* is shown as a white cartoon ribbon, *VanXYg* as a yellow cartoon ribbon. An extra domain can be observed, formed of two antiparallel β -sheets, in *VanXYg*.

The LdcB enzymes superpose less well with VanX (**Table 3.7**). The active site of VanX is much smaller and more constricted than that of *SpLdcB*, *BsLdcB* or *BaLdcB*, with the differences in active site shape likely arising from differing substrate specificities; a larger active site is required by LdcB enzymes to accommodate a ligand that is ~ 4 times the size as the D-Ala-D-Ala substrate of VanX.

Table 3.7: RMSD scores of the three LdcB enzymes (top) with VanXYg, VanXYc and VanX (side).

	<i>SpLdcB</i>	<i>BsLdcB</i>	<i>BaLdcB</i> *
VanXYg	RMSD: 1.6 Å # C α aligned: 146	RMSD: 2.0 Å # C α aligned: 167	RMSD: 1.5 Å # C α aligned: 154-156
VanXYc*	RMSD: 1.6 Å # C α aligned: 146-148	RMSD: 1.9 Å # C α aligned: 164-166	RMSD: 1.5 Å # C α aligned: 153-155
VanX*	RMSD: 2.3 Å # C α aligned: 97-100	RMSD: 1.9 Å # C α aligned: 112-116	RMSD: 1.9 Å # C α aligned: 110-119

* Multiple monomers are present within the asymmetric unit, RMSD values given are an average of the superimposition of all monomers and the number of aligned C α atoms is given as a range.

3.4.6 Comparing LdcB enzymes to other LD-carboxypeptidases

While the LdcB enzymes are the first of the LD-carboxypeptidases of the LAS family to be identified, several bacterial LD-carboxypeptidases have been characterised previously. These include LdcA (LD-carboxypeptidase A) from a range of organisms, including *E. coli* (Templin *et al.*, 1999), *Novosphingobium aromaticivorans* (Das *et al.*, 2013) and *Pseudomonas aeruginosa* (Korza and Bochtler, 2005), and the recently discovered Csd6 and Pgp2 enzymes, two homologous LD-carboxypeptidases which aid to maintain the helical shape of *H. pylori* (Sycuro *et al.*, 2013) and *C. jejuni* (Firdich *et al.*, 2014).

LdcA was first isolated from the cytoplasm of *E. coli* and is essential in maintaining cell viability during stationary phase growth (Templin *et al.*, 1999). It has been shown to cleave D-alanine from soluble tetrapeptides that are not incorporated into peptidoglycan (Templin *et al.*, 1999). This is in contrast to LdcB, which is found on the exterior of the cell and cleaves tetrapeptides found within peptidoglycan. A protein BLAST (Altschul *et al.*, 1990) comparison of the three LdcB homologues with various LdcA enzymes gives an average sequence identity of 67%, but with a coverage of only 3% (E-value 0.17). Unsurprisingly, given the lack of sequence homology, there is little structural conservation between these proteins. Comparing SpLdcB with either of the two LdcA structures present in the PDB (*P. aeruginosa* LdcA PDBid: 1ZRS or *Novosphingobium aromaticivorans* LdcA PDBid: 3G23), reveals that these enzymes adopt distinct folds, with core RMSDs of 3.8 Å over 56 aligned Cα atoms and 5.0 Å over only 65 Cαs obtained, respectively (**Figure 3.28**). LdcA utilises a Ser-His-Glu catalytic triad during catalysis and is therefore not a metalloprotease (Korza and Bochtler, 2005). Whilst it is clear that LdcA and LdcB share similar activities, they accomplish these functions by different mechanisms and are unrelated in structure and sequence.

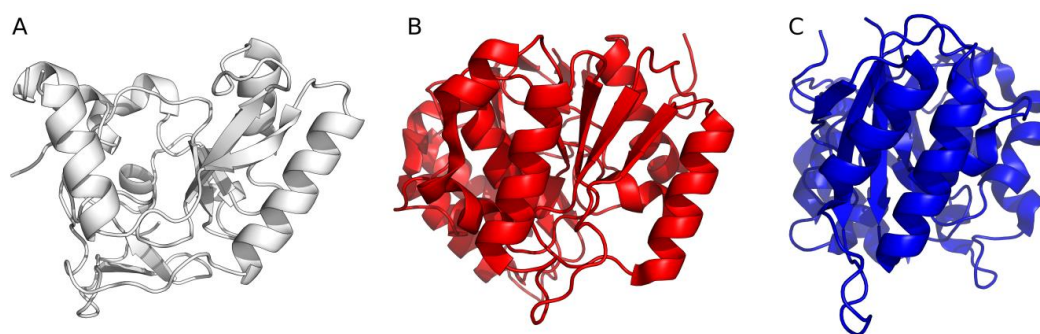


Figure 3.28: The structure of *SpLdcB* (A) compared to *LdcA* from *P. aeruginosa* (B) and *N. aromaticivorans* (C). All three proteins are shown as cartoon ribbons. The structure of *SpLdcB* shares no common structural motifs with either of the *LdcA* enzymes.

Equally, *LdcB* shares no sequence similarity with either *Csd6* or *Pgp2*. *Csd6* and *Pgp2* both belong to an LD-transpeptidase family that utilises a conserved Cys/His motif for catalysis (Biarrotte-Sorin *et al.*, 2006), and are therefore also distinct from the metalloprotease *LdcB*. Interestingly, *Csd6* has been shown to only hydrolyse D-Ala from monomeric stem peptides whereas *LdcB* is active against both monomeric and cross-linked (tetra-tetra) stem peptides (Sycuro *et al.*, 2013; Hoyland *et al.*, 2014), but the absence of any structure or sequence similarity between these enzymes precludes useful comparison.

3.5 Summary

This chapter describes the crystallisation and structure determination of proteins encoded by the *ldcB* gene in *S. pneumoniae* and *B. subtilis*. The crystal structures allowed key features of these enzymes to be identified and comparisons made to related enzymes.

Initial crystallisation trials of *SpLdcB* produced rod-shaped crystals that diffracted to 3.6 Å. These crystals were not suitable for high resolution structural analysis. The addition of excess zinc following purification was shown to greatly enhance the quality of crystals produced.

Diffraction data were collected from the improved zinc-bound Set-Met substituted *SpLdcB* crystals and the structure solved by Se-Met SAD. In addition, *BsLdcB* was crystallised and native diffraction data were collected. The *BsLdcB* structure was solved by molecular replacement with that of *SpLdcB* as the search model.

The structure of LdcB reveals a single-domain protein comprised of two-subdomains separated by a V-shaped cleft, at the bottom of which is a Zn^{2+} , coordinated by two histidines and an aspartate that extend from a four-stranded, antiparallel β -sheet, a characteristic feature of LAS family of peptidases (Bochtler *et al.*, 2004). In addition, the presence of an E-x-x-H motif, in place of H-x-H, places LdcB enzymes in a subset of LAS enzymes alongside VanX, VanY, VanXYg and VanXYc. The similarities in structure of LdcB with VanY, VanXYg and VanXYc, in addition to their cleaving of substrates that differ in length by only one amino acid, suggests that these enzymes bind and catalyse the hydrolysis of their substrates in similar manners. Understanding the structure and function of LdcB would therefore provide key insights into not only LD-carboxypeptidases, but potentially vancomycin resistance also, a subject of great interest in this era of increasing antibiotic resistance.

The results presented herein provide the first structures of zinc-utilising LD-carboxypeptidases. Residues important in the binding of peptidoglycan stem-peptides are highlighted through the appearance of conserved residues in VanXYg and VanXYc, though questions on how these enzymes differentiate between pentapeptide and tetrapeptide substrates remain to be answered.

Chapter 4:

**The structure of *Sp*LdcB complexed with
MurNAc-Ala-D- γ -Gln-Lys(D-Asn) & the biochemical
characterisation of LdcB enzymes**

4.1 Introduction

LdcB has been implicated previously in cleaving D-alanine from cell wall tetrapeptides; the cell walls of *ldcB* null mutants in *S. pneumoniae*, *L. lactis* and *B. subtilis* showed an increase in the amounts of tetrapeptides accompanied by a decrease in tripeptide content (Courtin *et al.*, 2006; Barendt *et al.*, 2011; Hoyland *et al.*, 2014). However, no direct observation of the catalytic activity of LdcB has been reported. The crystal structures of *Sp*LdcB and *Bs*LdcB that were described in Chapter 3 and by Hoyland *et al.*, (2014), and that of *Ba*LdcB, which was solved by the CSGID and that was also reported by Hoyland *et al* (2014), require complementing with an LdcB structure in complex with a peptidoglycan fragment in order to understand substrate specificity and the catalytic mechanism employed at a molecular level by this enzyme family. Such knowledge is also likely to provide key insights on molecular recognition and catalysis in the wider LAS family of peptidases. In this chapter the crystal structure of LdcB with a bound product mimic is described (**Section 4.2**), the catalytic properties of LdcB are investigated (**Sections 4.3 & 4.4**), and an enzymatic mechanism for LdcB is proposed (**Section 4.5**).

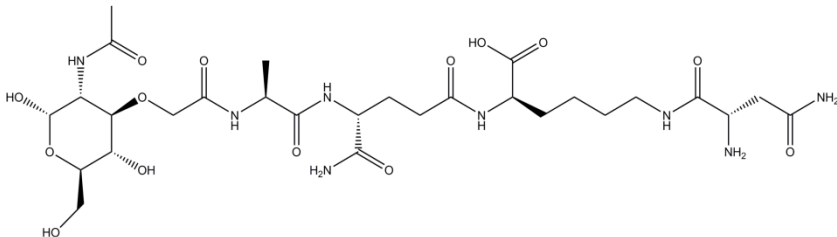
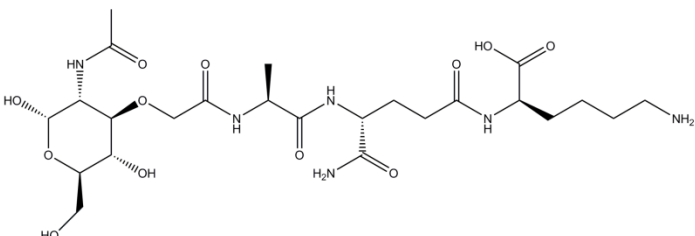
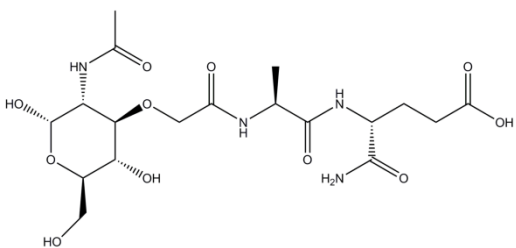
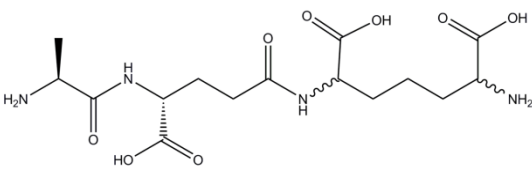
4.2 Co-crystallisation

4.2.1 Selection of ligands for co-crystallisation studies

Although the crystal structures of *Sp*LdcB and *Bs*LdcB revealed key structural features of the enzymes, the enzymes in complex with a ligand would allow detailed analysis of the residues that line the active site and allow a determination of how the enzyme differentiates between L and D peptides. Since wildtype LdcB would be catalytically active against cell wall tetrapeptides, their use in co-crystallisation experiments was precluded. A further problem is the general unavailability of many of the precursors required for cell wall synthesis on a scale that supports structural studies. In instances where the target ligand is unavailable, is unstable or is degraded by the enzyme, a common method employed to trap enzyme:substrate complexes is with non-hydrolysable substrate or product analogues (Oswald *et al.*, 2008). For example, AMP-PNP or ADP have been substituted successfully for ATP in co-crystallisation studies on kinases (Ivey *et al.*, 2004; Simanshu *et al.*, 2005). In this instance, four

peptidoglycan mimics were found to be commercially available from InvivoGen; M-Tri-Lys, M-Tri-Lys(D-Asn), MDP and Tri-DAP (**Table 4.1**).

Table 4.1: The chemical structures of commercially available cell wall fragments. All fragments were purchased from InvivoGen (<http://www.invivogen.com>)

Name	M _r	Structure
M-Tri-Lys(D-Asn)	734.75	
M-Tri-Lys	620.65	
MDP	492.5	
Tri-DAP	390.39	

4.2.2 Crystallisation trials of *Sp*LdcB with a range of products and substrates

To discover the molecular basis of substrate specificity in LdcB enzymes, the crystal structure of *Sp*LdcB in complex with a range of commercially-available ligands was sought (**Table 4.1**). Native *Sp*LdcB was purified and concentrated to 15 mg/ml and a 4:1 molar excess of zinc was added (**Section 2.2**). The target ligand was added to a final concentration of 2 mM from concentrated stocks to avoid excessive dilution of the

protein. Crystallisation trials were conducted in 96-well sitting-drop vapour diffusion trays, utilising a Mosquito (TTP labtech) crystallisation robot and a range of commercially available screens (Molecular Dimensions, Qiagen, Hampton Research), with crystals growing over a period of 2 – 7 days. The crystallisation conditions were not optimised due to the limited availability of the ligands. Crystals were obtained for each of the four ligands under the different conditions, but with some overlap; all screens produced crystals from PACT C8, whereas only MDP produced crystals in Structure G8 (for a full list of conditions, see **Appendix 4**). Crystals (**Figure 4.1**) were harvested directly from the 96-well screen plates, cryo-protected in paratone-N oil and flash cooled in liquid nitrogen. Due to the proximity of a synchrotron trip, crystals were not tested in-house prior to diffraction data collection at the Diamond Light Source on beamline IO2; each dataset was collected at 0.92 Å with 0.1° oscillations and 0.1 s exposures per image. The total number of images collected was based upon the predicted space group of each individual crystal to ensure that complete data sets were collected.

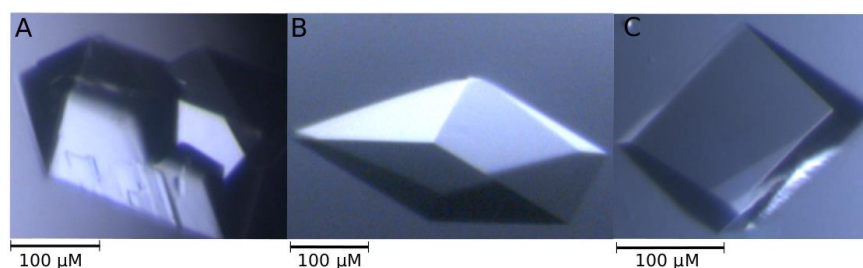


Figure 4.1: Crystals of *SpLdcB* co-crystallised with (A) Tri-DAP; (B) M-Tri-Lys(D-Asn); (C) M-Tri-Lys

Each diffraction dataset was integrated with XDS (Kabsch, 2010) and reduced with SCALA (Evans, 2005). Diffraction data collection statistics for each crystal grown in the presence of a unique ligand are summarised in **Table 4.2**. Interestingly, it was noted that dataset 3 shares an almost identical unit cell with native *SpLdcB*, despite belonging to different space groups; native *SpLdcB* was found to be *I*222, whereas dataset 3 is *P*2₁2₁2. Comparing the packing of the molecules in the unit cell showed that both cells were equivalent, suggesting that the difference in space group was due to the presence of two molecules in the asymmetric unit dataset 3 (*P*2₁2₁2) and one

molecule in the native *SpLdcB* dataset (*I*222); non-crystallographic symmetry operators are present in dataset 3, whereas the native *SpLdcB* unit cell is constructed from crystallographic symmetry operators. An example of an identical situation has previously been highlighted on the ccp4 bulletin board (personal communication), using cellobiohydrolase I from *Trichoderma reesei* (Stahlberg *et al.*, 1996) as a case study.

The structure of *SpLdcB* was used as the molecular replacement model in PHASER (McCoy *et al.*, 2007) to solve the structures of all four datasets. Following one round of rigid body refinement in REFMAC5 (Winn *et al.*, 2011), difference electron density maps were inspected to determine the presence, or absence, of ligand in the vicinity of the active site zinc ion. Only in dataset (1), in which PHASER had positioned 5 chains in the asymmetric unit, was positive electron density found that corresponded to M-Tri-Lys(D-Asn) in just one of the protein chains.

Table 4.2: Summary of representative X-ray data collection statistics for *SpLdcB* with ligands.

Dataset	1	2	3	4
Co-crystallised ligand	M-Tri-Lys(D-Asn)	M-Tri-Lys	MDP	Tri-DAP
Condition	PACT C8	Morpheus D8	Morpheus E8	PACT C8
Space Group	C2	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 2 ₁ 2 ₁ 2 ₁
Unit cell dimensions				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	346.0, 42.5, 79.3	57.9, 139.1, 45.2	48.2, 58.4, 137.1	48.5, 59.2, 138.1
α , β , γ (°)	90, 93.1, 90	90, 90, 90	90, 90, 90	90, 90, 90
Wavelength (Å)	0.92	0.92	0.92	0.92
Resolution (Å)	47.8 - 2.80 (2.95 - 2.80)*	53.43 - 2.30 (2.42 - 2.30)	45.71 - 2.15 (2.22 - 2.15)	46.05 - 1.90 (1.94 - 1.90)
Multiplicity	3.4 (3.4)	3.2 (3.1)	3.2 (3.4)	7.0 (5.3)
<i>R</i> _{merge} ^a	0.134 (0.635)	0.047 (0.325)	0.031 (0.356)	0.144 (0.319)
<i>I</i> / σ <i>I</i>	7.8 (1.9)	13.7 (3.4)	16.7 (2.2)	8.9 (3.3)
Completeness (%)	98.3 (98.3)	98.1 (97.7)	98.6 (97.5)	100 (100)

* Highest resolution shell shown in parentheses.

^a $R_{\text{merge}} = \sum_h \sum_i |I_i(h) - \langle I(h) \rangle| / \sum_h \sum_i I_i(h)$, where $I_i(h)$ and $\langle I(h) \rangle$ are the *i*th and mean measurement of the intensity of reflection *h*.

The model was refined in COOT through stages of manual (re)building interspaced with restrained refinement in PHENIX.REFINE (Afonine *et al.*, 2012) with automatically generated NCS restraints, until convergence. TLS groups were identified through the use of the TLS server (Painter and Merritt, 2006) and 4 TLS groups were used to refine chains C and D, 2 groups were used for chains B and E, and 5 groups were used for

chain A. The MurNAc-L-Ala-D-γ-Gln-L-Lys-(D-Asn) ligand was built using JLigand (Lebedev *et al.*, 2012) to output structure (.pdb) and restraint (.cif) files, and the ligand was subsequently positioned manually in the electron density in COOT. Restrained refinement in PHENIX.REFINE (Afonine *et al.*, 2012) was used to optimise the position of the ligand.

The final model of ligand-bound *SpLdcB* consisted of 5 protein chains, in which all cloned residues (56 - 237) except the terminal residue, D238, were present. In addition, histidine (-2) and methionine (-1) - contributed by the introduction of an N-terminal His-tag in pET28a - could be built in chains A, B, D and E. Density for chains A and B also permitted the building of serine (-3). The Ramachandran plot of the final structure revealed that 96.5 % of residues lay within the favoured regions, with 2.9 % found in allowed regions and just 5 Ramachandran outliers (**Figure 4.2**).

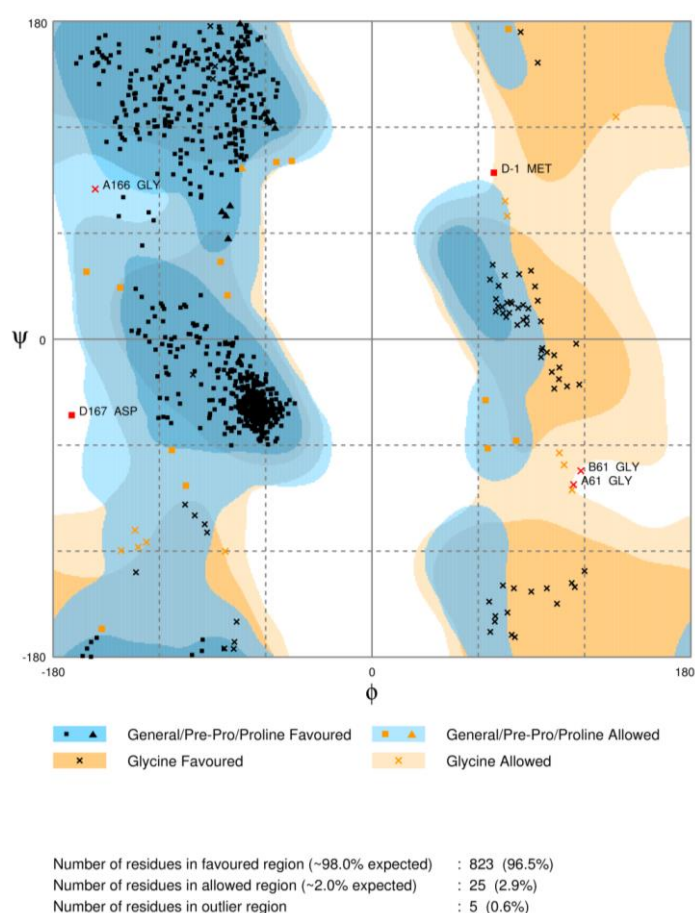


Figure 4.2: Ramachandran plot of the *SpLdcB*-ligand structure. Black shapes represent residues in favoured regions, orange shapes represent those residues in allowed regions. This plot was generated using the RAMPAGE webserver (Lovell *et al.*, 2003) (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>)

The electron density of each Ramachandran outlier was inspected, with each amino acid fitting well into the observed density. With the exception of chain D, Met-1, which is at the N-terminus of the chain, the outliers were all internal residues that neither contributed to crystal contacts nor formed part of the active site. A summary of the final refinement statistics can be found in **Table 4.3**.

Table 4.3: Summary of X-ray data refinement statistics for *SpLdcB* + M-Tri-Lys(D-Asn).

PDBid	<i>SpLdcB</i> -ligand
	4OXD
Refinement	
Resolution (Å)	47.8-2.80
No. Reflections	47327 (1979)
$R_{\text{work}}/R_{\text{free}}^a$	0.273/0.335
No. atoms	
Protein	7014
# proteins/AU	5
Ion/Ligand	21/51
Water	111
B factors (Å ²)	
Protein	44.3
Ion/Ligand	38.7/69.7
Water	28.7
R.m.s deviations	
Bond lengths (Å)	0.005
Bond angles (°)	0.84
Ramachandran	
favoured (%)	95.8
allowed (%)	99.6

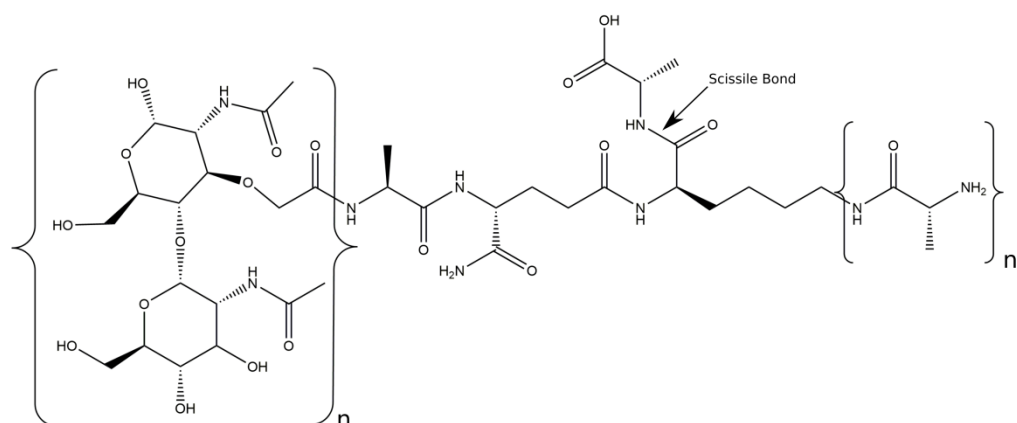
* Highest resolution shell shown in parentheses.

^a $R_{\text{work}}/R_{\text{free}} = \sum |F_{\text{p}}^{\text{obs}} - F_{\text{p}}^{\text{calc}}| / \sum F_{\text{p}}^{\text{obs}}$, where $F_{\text{p}}^{\text{obs}}$ and $F_{\text{p}}^{\text{calc}}$ are the observed and calculated structure factor amplitudes, respectively.

4.2.3 The structure of *SpLdcB* with M-Tri-Lys(D-Asn) bound to the active site

The structure of *SpLdcB* complexed with MurNAc-L-Ala-D-γ-Gln-L-Lys-(D-Asn) was solved. This ligand is a close relative of the true reaction product (**Figure 4.3**, overleaf) and differs in two ways. First, the ligand lacks GlcNAc, the second sugar of the disaccharide that is the glycan repeating unit in PG. Second, the terminal D-Asn group is covalently attached by an isopeptide bond between the main chain carbonyl of the terminal D-Asn and the side chain N_ε of the penultimate lysine. In cross-links between peptide stems of *S. pneumoniae*, the side chain N_ε of the penultimate lysine is linked by a peptide bond to the main chain carbonyl of a D-Ala. Hence the L-Lys—(D-Asn) linkage is a faithful mimic of the natural L-Lys—(D-Ala), albeit the chemistry of the side chains of the terminal D-amino acids clearly differ.

A



B

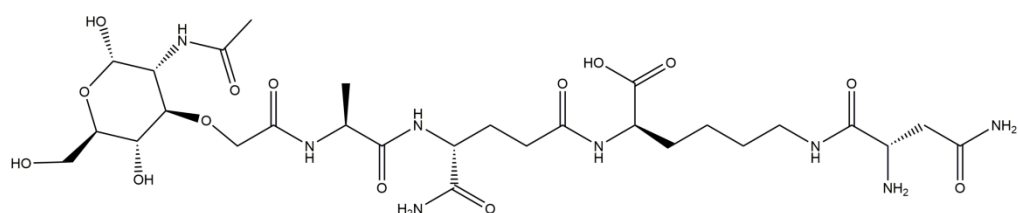


Figure 4.3: Schematic representations of the chemical structures of the true LdcB reaction product (A) and the product mimic (B). The product mimic lacks the repeating GlcNAc-MurNAc sugars of the true substrate. Additionally, the D-alanine that is covalently attached to the N_ε of the lysine residue has been replaced by D-asparagine. Both crosslink types are found in many bacterial species, and thus the Lys-(D-Asn) linkage is faithful to the crosslinks found in nature.

Analysis of the interfaces between the 5 protein chains in PISA (Krissinel and Henrick, 2007) indicated that none of them represents stable multimeric forms (all buried surface areas < 5% total protein surface). The five molecules are highly superimposable (**Table 4.4**, overleaf), but some key differences in the active sites remain. For instance, in each chain B, C and E, the zinc is co-ordinated by a likely nucleophilic water to form tetrahedral co-ordination geometry. In chain D, the fourth zinc ligand is not evident, perhaps due to local disorder in the active site or the limiting resolution of the diffraction data. In the remaining molecule, chain A, residual electron density was retained throughout refinement into which MurNAc—L-Ala—D-γ-Gln—L-Lys—(D-Asn) could be built and refined unambiguously (**Figure 4.4**, overleaf).

Table 4.4: RMSD values between Ligand-bound *SpLdcB* chains (4OXD)

	4OXD Chain A	4OXD Chain B	4OXD Chain C	4OXD Chain D	4OXD Chain E
4OXD Chain A	-	0.63	0.64	0.76	1.03
4OXD Chain B	0.63	-	0.55	0.65	0.86
4OXD Chain C	0.64	0.55	-	0.56	0.64
4OXD Chain D	0.76	0.65	0.56	-	0.61
4OXD Chain E	1.03	0.86	0.64	0.61	-

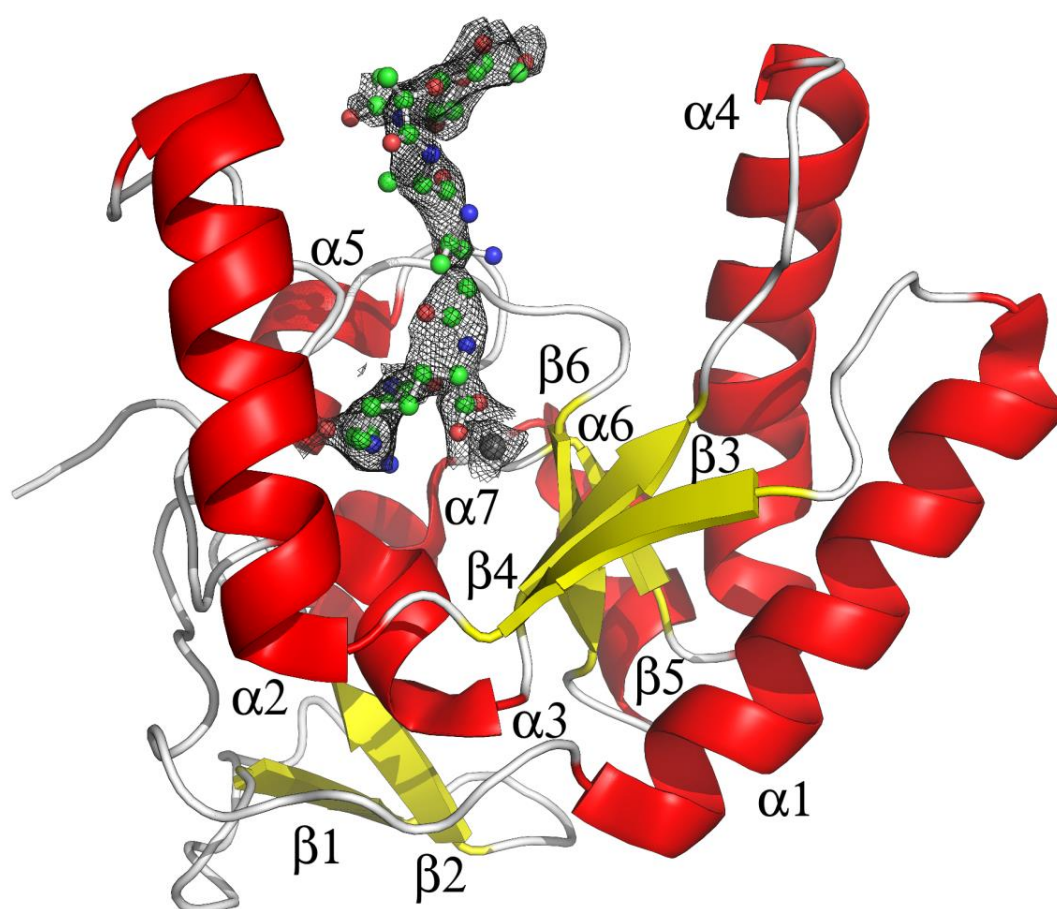


Figure 4.4: Cartoon representation of ligand-bound *SpLdcB*. The MurNAc—Ala—D-γ-Gln—L-Lys—(D-Asn) is drawn as a ball-and-stick model with Refmac-weighted $2F_{\text{obs}} - F_{\text{calc}}$ electron density displayed at a contour level of 1σ in black. α -helices are shown in red, β -strands in yellow and loops in white. The zinc ion is shown as a grey sphere.

The structure of the ligand-bound form of *SpLdcB* is entirely consistent with its LD-carboxypeptidase function. First, the MurNAc moiety, which is not involved in catalysis, is situated at the very top of the inter-subdomain cavity and makes little

contact to the protein. This part of the ligand has the highest B-factors (mean of 93 Å³, vs a mean for the intact ligand of 83 Å³), and thus the weakest electron density. The MurNAc O1 and O4 atoms, which define the location of the flanking, β1-4-linked GlcNAc sugars in PG, project away from the protein surface so that the glycan chain would be accommodated by the protein without steric clashes. The oxygen of the N-acetyl group is stabilised against the side chain of Met202 by van der Waals' forces (**Figure 4.5**), however this residue, and therefore interaction, is not conserved across LdcB homologues.

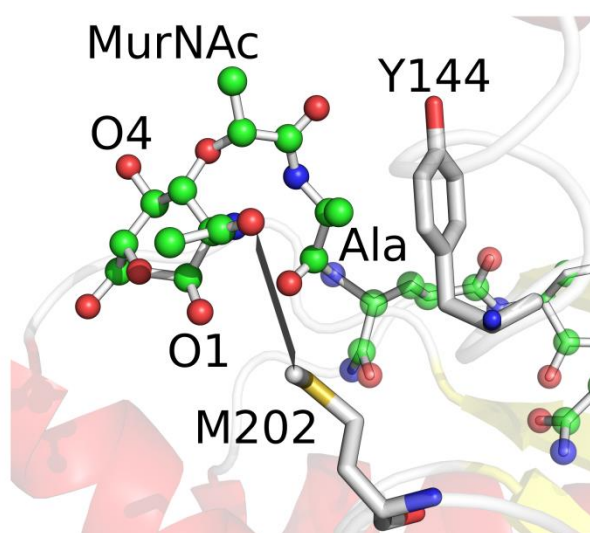


Figure 4.5: Interactions between the M-Tri-Lys(D-Asn) substrate and *SpLdcB* at subsite *S*₃. The MurNAc sugar of the M-Tri-Lys(D-Asn) ligand forms Van der Waals contacts with M202 (silver line), whilst the L-stereoisomer of alanine is selected for by Tyr144. The M-Tri-Lys(D-Asn) ligand is shown as ball and stick representation.

The C α and C β atoms of L-alanine at position 1 of the peptide pack against the face of Tyr144 to discriminate against the D-stereoisomer in this subsite, *S*₃. L-alanine is favoured by contacts to Tyr144 and Met202, which discriminate against all L-amino acids bulkier than alanine (**Figure 4.5**). The NH₂ group of the D- γ -Gln of position 2 is located in subsite *S*₂, within hydrogen bonding distance of Glu204 (**Figure 4.6**, overleaf), an invariant amino acid in LdcB sequences, and stereoselectivity at this position is maintained because L- γ -glutamine would clash with Tyr144.

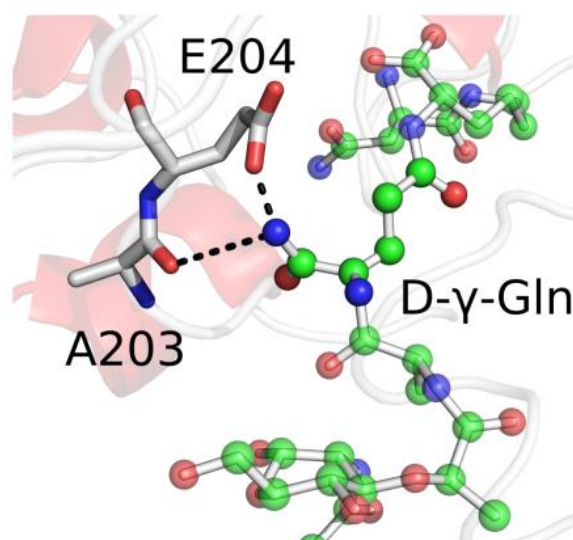


Figure 4.6: Interactions between the M-Tri-Lys(D-Asn) substrate and *SpLdcB* at subsite S_2 . E204 and A203 coordinate the D-γ-Gln. These contacts would not occur with commonly occurring amino acids, providing selectivity for D-γ-Gln. Hydrogen bonds are shown as black, dashed lines. The M-Tri-Lys(D-Asn) ligand is shown as ball and stick representation.

The γ-glutamine is selected for by specific interactions with the main chain carbonyl of Ala203 and the side chain of Glu204, whilst ensuring the scissile peptide bond is in the catalytic position. These key aspects of molecular recognition could not occur with commonly occurring amino acids. It is possible that canonical glutamine could be accommodated at this position, due to the absence of contacts to one face of the peptide, providing scope for the conformational changes needed to accommodate glutamine. The substitution of glutamine for γ-glutamine at this site would, however, only be possible in the absence of the glycan strands of the natural substrate, due to a 'shortening' of the chain (**Figure 4.7**, overleaf).

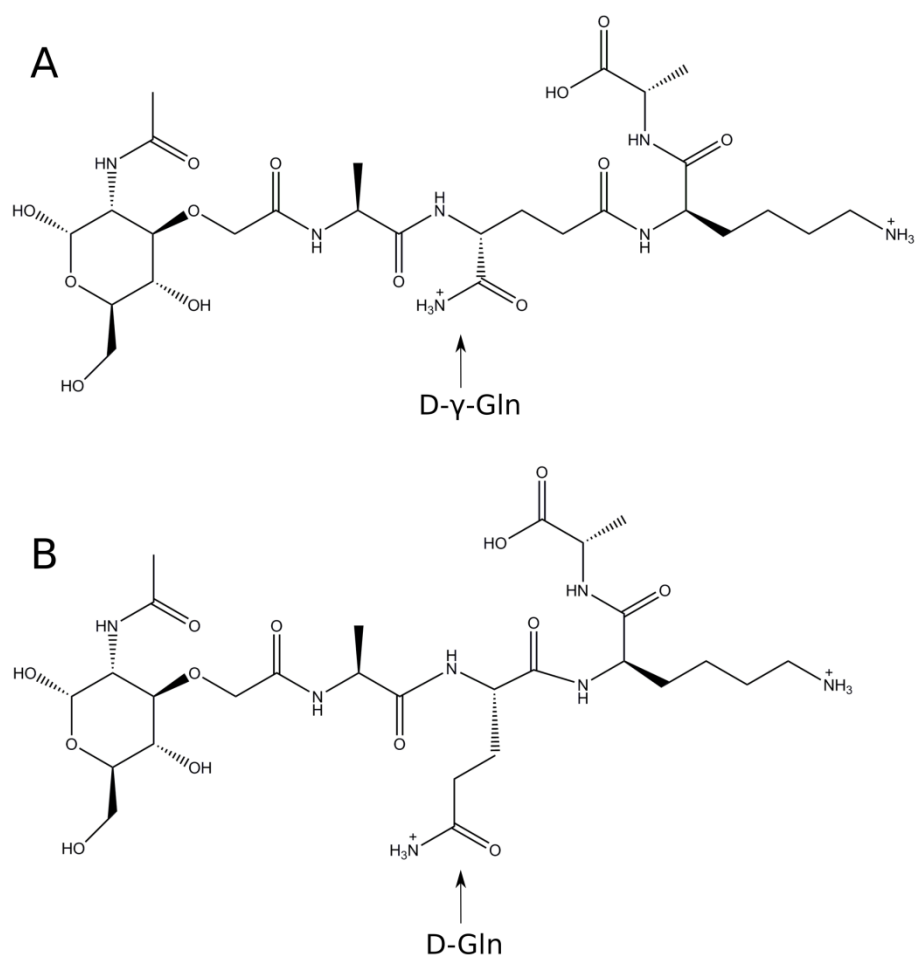


Figure 4.7: The structure of MurNAc-tetrapeptide (A) compared to a mimic in which D- γ -Gln is substituted by D-Gln (B). The substitution results in a shortening of the chain by ~ 3 Å.

Position 3 of the peptide is occupied by L-lysine. D-lysine could not be accommodated at this position because of the presence nearby of Glu107. As expected of a reaction product complex, the lysine carboxylate is located adjacent to the active site zinc in subsite S_1 (**Figure 4.8**, overleaf). One of the carboxylic oxygens completes the zinc tetrahedral geometry with His153/Asp160/His207 and the second oxygen forms a salt bridge with Arg120, which is either arginine or lysine in LdcB orthologues. The L-lysine carboxylate thus represents a post-hydrolysis state, placed appropriately for the peptide bond of the substrate to be hydrolysed by the water molecule that fulfils the fourth zinc ligand in the absence of substrate.

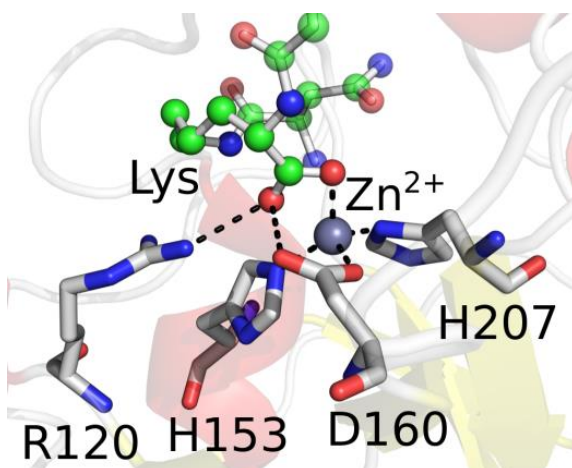


Figure 4.8: Interactions between the M-Tri-Lys(D-Asn) substrate and *SpLdcB* at subsite S_1 . The ligand displaces the solvent molecule from the zinc coordination sphere and becomes the 4th ligand. The second oxygen of the lysine carboxyl group interacts with Arg120 and Asp160. Hydrogen bonds are shown as black, dashed lines. The M-Tri-Lys(D-Asn) ligand is shown as ball and stick representation.

The sidechain of the L-lysine twists back against the rest of the peptide and projects into the protein core to place the terminal D-Asn in subsite S_1' against the main chain atoms of Ala146 and the side chains of Tyr191, Tyr201, (**Figure 4.9**) The interactions between the protein and the ligand are summarised schematically in **Figure 4.10** (overleaf).

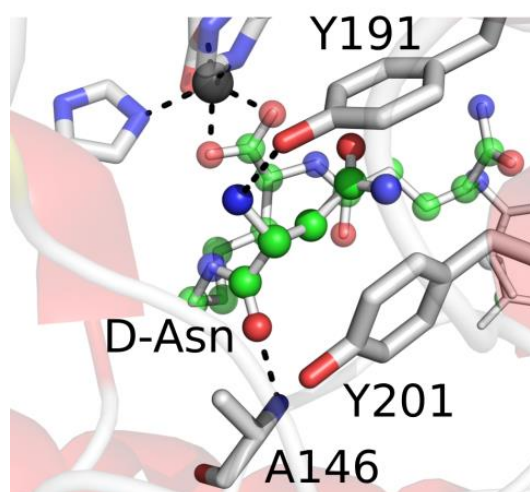


Figure 4.9: Interactions between the M-Tri-Lys(D-Asn) substrate and *SpLdcB* at subsite S_1' . The D-Asn, linked to the L-Lys at position 3 by an iso-peptide bond, interacts with Ala146, Tyr201 and Tyr191. Hydrogen bonds are shown as black, dashed lines. The M-Tri-Lys(D-Asn) ligand is shown as ball and stick representation.

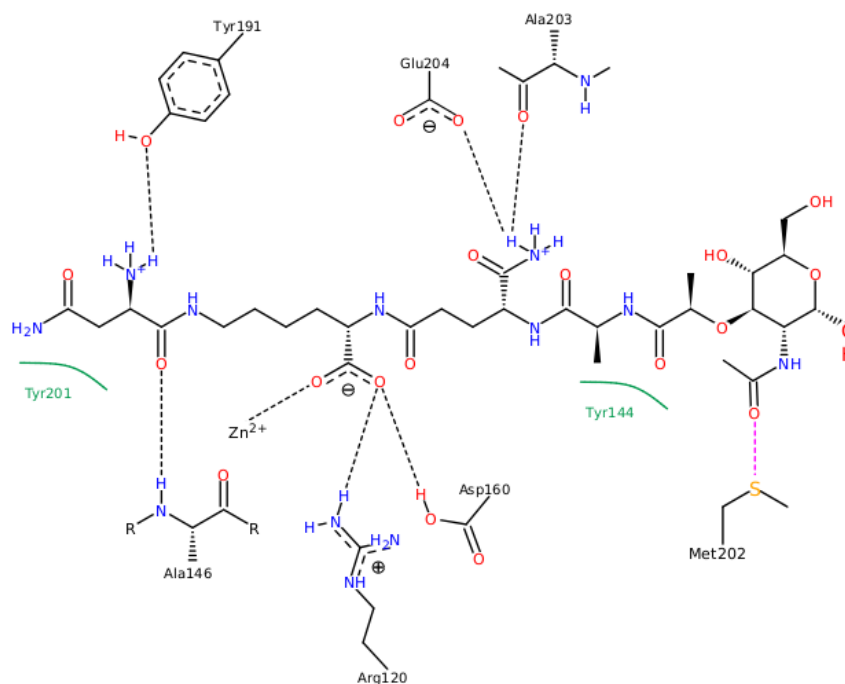


Figure 4.10: Schematic summary of the interactions made between *SpLdcB* and the product mimic, **M-Tri-Lys(D-Asn).** Hydrogen bonds are shown with dotted black lines, hydrophobic interactions shown by solid green lines and Van der Waals forces by a purple dotted line. The figure was made with Poseview (Stierand *et al.*, 2006; Stierand and Rarey, 2010) (<http://poseview.zbh.uni-hamburg.de/>)

Excluding metal binding residues and the catalytic arginine and glutamate, all the residues that contact the tetrapeptide ligand are strictly conserved, with the exception of M202, which only interacts with the MurNAc sugar by van der Waal's interactions; Ala203, which interacts via a carbonyl oxygen and thus can be substituted by all amino acids; Y144, which is functionally conserved as either a tyrosine or phenylalanine; Ala146, which is highly conserved and again, interacts via a mainchain atom; Y191, which is functionally conserved as either a tyrosine or phenylalanine; and Y201, which is functionally conserved as either a tyrosine or isoleucine. In addition, there are several stretches of amino acids that are either strictly or highly conserved that are not found near the active site (**Figure 4.11**). One stretch of conserved residues (151-160 in *SpLdcB*) are found to line the lower part of the active site and may prevent LdcB from binding larger ligands. However, their conservation in the VanY family enzymes, which bind pentapeptides, suggest that these ligands would have to be larger than a pentapeptide to be excluded from the active site.

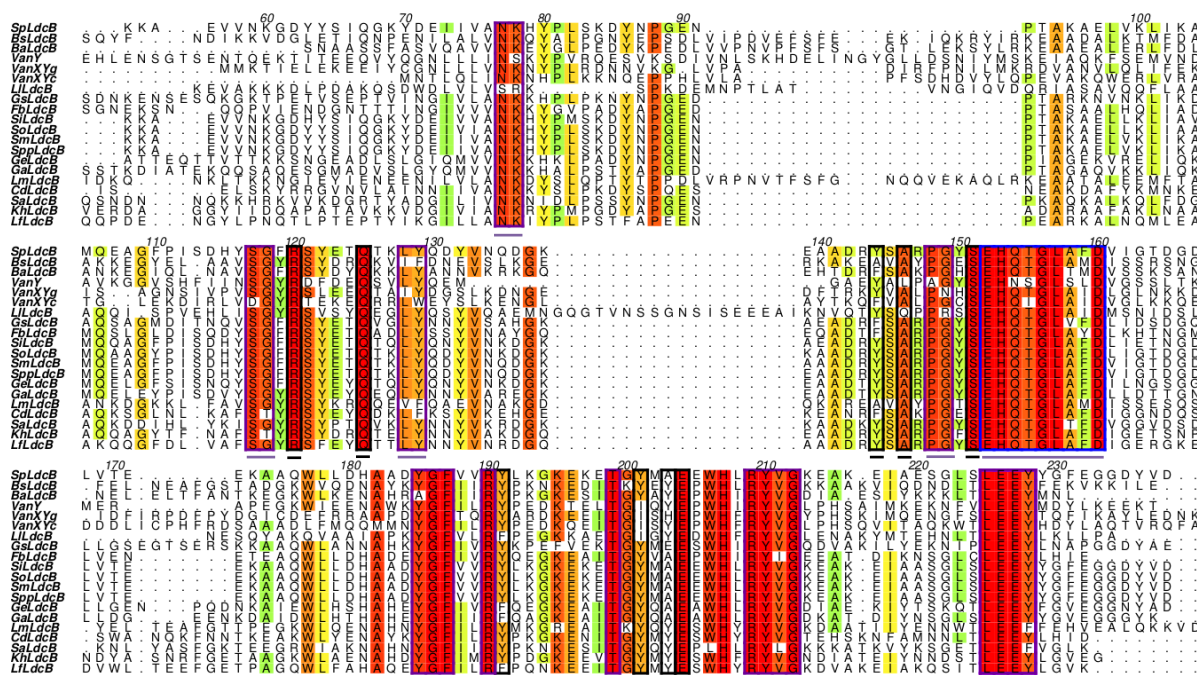


Figure 4.11: Sequence alignment showing residue conservation in a range of LdcB and VanY-like enzymes. Conserved amino acids involved in ligand binding are shown in black boxes with black underline. A blue box and blue underline indicates conserved amino acids located at the bottom of the active site. Purple boxes with purple underline indicate conserved amino acids located at other sites on the protein. Residues are numbered in reference to the SpLdcB sequence. *Spp* = *S. pseudopneumoniae*, *Sm* = *S. mitis*, *Si* = *S. infantis*, *So* = *S. oralis*, *Gs* = *G. sanguinis*, *Lf* = *L. fusiformis*, *Ga* = *G. adiacens*, *Ge* = *G. elegans*, *Fb* = *Fermicute bacterium M10-2*, *Kh* = *K. huakuii*, *Ll* = *L. lactis*, *Cd* = *C. difficile*, *Sa* = *S. aureus* and *Lm* = *L. monocytogenes*).

In the true enzyme-substrate complex, subsite S_1' would be occupied by the scissile D-alanine; the structure of VanXYg solved in complex with a phosphinate transition state mimic (PDBid: 4muq, (Meziane-Cherif *et al.*, 2014) supports this hypothesis. The phosphinate moiety is a transition state mimic of peptide bond hydrolysis because the reaction proceeds *via* an unstable oxyanion tetrahedral intermediate that the chemically-inert phosphinate mimics. VanXYg superimposes on SpLdcB with an RMSD of 1.6 Å on 144 matched Cαs, and the alanyl carboxylate moiety of the phosphinate is situated in an analogous position to that of the L-Lys—(D-Asn) peptide linkage observed in SpLdcB (Figure 4.12, overleaf). The phosphate atoms of the phosphinate transition state mimic superimpose almost identically with the phosphate anion found adjacent to the zinc ion in the structure of BsLdcB (Figure 3.22, Section 3.4). In addition to Tyr191 and Tyr201, which interact with D-Asn, Ser151 and Gln125 (strictly conserved in LdcB sequences) aid to determine specificity for the terminal D-alanine by hydrogen bonding with the carboxylic oxygens of D-Ala (Figure 4.12, overleaf); the carboxylic oxygens of L-Ala would be orientated in a position too distant for hydrogen

bonding to these amino acids. Tyr201 discriminates against any amino acid bulkier than alanine at the S₁' subsite, though the D-Asn of the product mimic is permitted because it is linked to L-lysine by an isopeptide bond rather than by a canonical peptide bond. The amide nitrogen of Ala146 also participates in stereoselectivity, but as this is a main chain atom it is perhaps no surprise that sequence conservation is not maintained at this position.

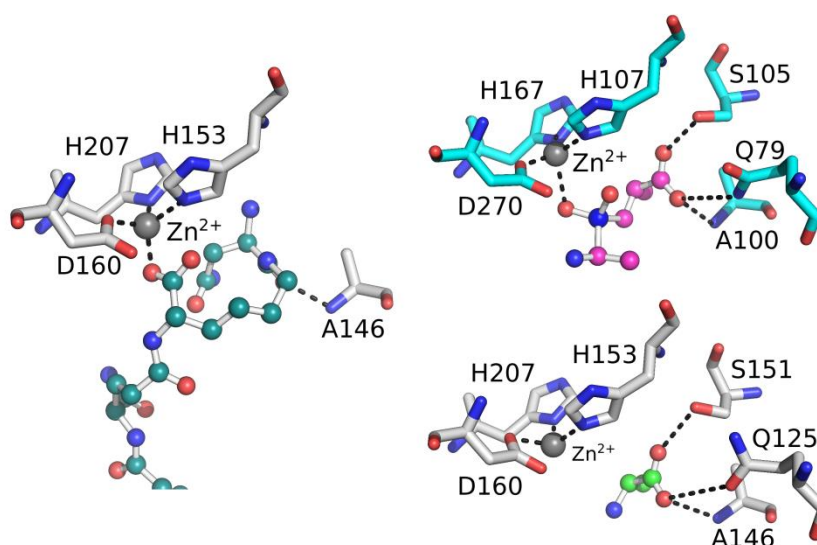


Figure 4.12: A comparison of the ligand-bound active sites carboxypeptidases. (A) The active site of *SpLdcB* (white) with bound MurNac—L-Ala-D-γ-Gln—L-Lys—(D-Asn) (turquoise carbons) is compared to that of the active site of VanXYg (blue) containing a phosphinate transition state analog (pink carbons) of D-Ala—D-Ala (PDBid: 4MUQ). The phosphinate in VanXYg coordinates the active site zinc in a similar manner to that of the lysyl carboxylate in the MurNac—L-Ala-D-γ-Gln—L-Lys—(D-Asn) bound to *SpLdcB*. The D-alanyl moiety of the phosphinate ligand in VanXYg, which occupies subsite S1', is matched by the D-alanine (green carbons) in Apo *SpLdcB* (Section 3.4.2) (white) and the D-Asn of the MurNac—L-Ala-D-γ-Gln—L-Lys—(D-Asn) ligand. The zinc ions are shown as grey spheres.

Although the structure of *SpLdcB* in complex with M-Tri-Lys(D-Asn) revealed important aspects of ligand binding, peptidoglycan cross-links between peptide stems could not be accommodated in the observed conformation of the ligand. Rotation of the sidechain torsion angles of the L-lysine into an extended, thermodynamically-favoured conformation results in the side chain passing beyond Arg120, Leu128 and Tyr132 to project beyond the surface of the protein (Figure 4.13, overleaf).

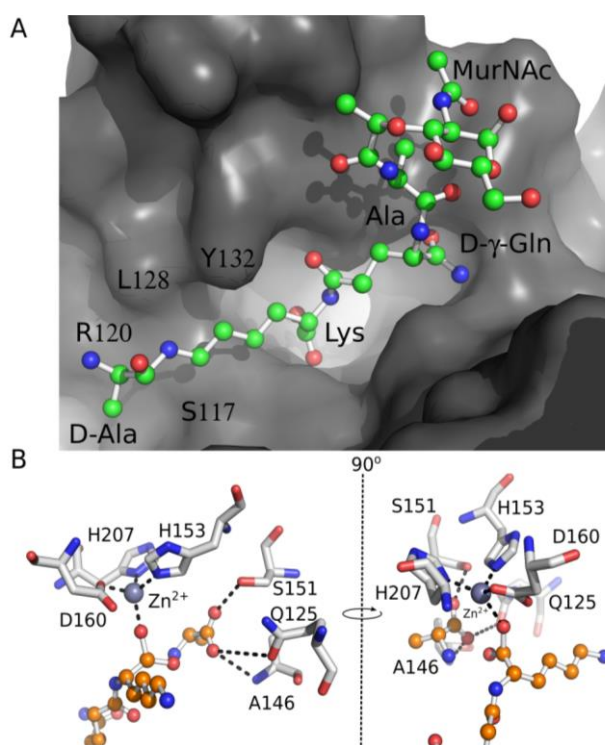


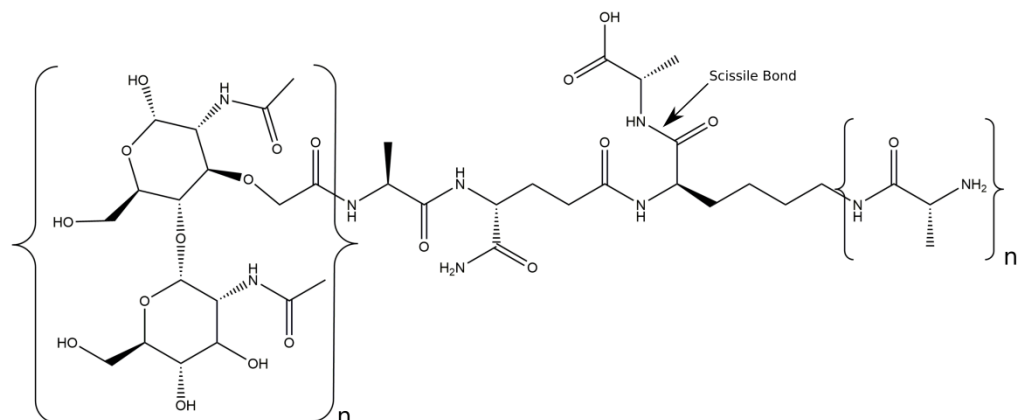
Figure 4.13: Proposed path of cross-linked cell wall peptides.

(A) The terminal D-Asn of the ligand when bound to *SpLdcB* occupies the S_1' subsite rather than projecting to the surface of the protein. Rotation of the lysyl's torsion angles presents the terminal D-asparagine to the surface, which has been modelled here as D-Ala to illustrate the structure of the L-Lys-D-Ala cross-link observed between stem peptides. The path of the lysine is flanked by two highly conserved amino acids in LdcB sequences, Arg120 and Ser117, in addition to the functionally conserved Leu128. Tyr132 also flanks the lysine, but this amino acid is not conserved between sequences.

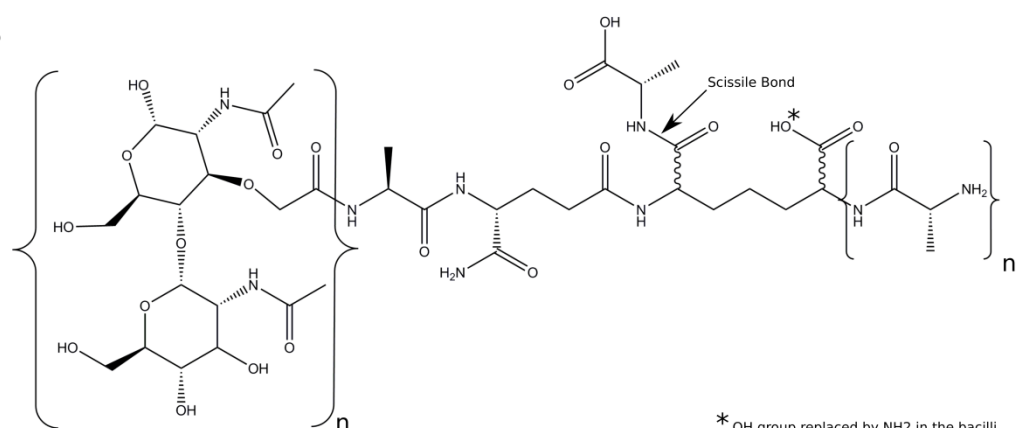
(B) The proposed positions of the two, post-hydrolysis ligands (orange carbons) in the active site of *SpLdcB* (white). The D-alanine from the *SpLdcB* structure occupies the S_1' site of the protein, whilst the lysine side chain has been rotated and modelled in an extended conformation. The altered conformation of the lysine presents the N_ζ to the surface of the protein, in an ideal position to accommodate peptide cross-links. The zinc ions are shown as grey spheres.

In this conformation, cross-linked peptide stems attached to the terminal N_ζ would not be occluded by the body of the protein, consistent with the reported ability of LdcB to cleave cross-linked peptides (Hoyland *et al.*, 2014). In some bacteria, the lysine is replaced with meso-diaminopimelic acid, a carboxy derivative of lysine that is amidated in bacilli (Figure 4.14, overleaf). Whilst the aminocarboxy derivative could be stabilised in the observed conformation by a similar set of specific interactions, it would also occlude cross-linked peptide stems. Diaminopimelic acid is thus likely to follow the same path predicted for lysine, by forming a simple, extended conformation to project into solvent.

A



B



* OH group replaced by NH₂ in the bacilli

Figure 4.14: L-Lysine (A) vs. meso- diaminopimelic acid (B) containing stem peptides. Meso-DAP is an epsilon-carboxy derivative of lysine, and thus the two stem peptides are of the same length. The site of amidation in the meso-DAP peptide is shown by an asterisk.

Comparing the topology of the ligand bound to SpLdcB to that of cell wall fragments bound to other cell wall binding proteins (i.e. PDBid 2WWD, 2APH, 2CB3) reveals that the ligand is flexible enough to follow multiple paths through the various active sites of the many peptidoglycan binding proteins (cell wall synthases, hydrolases, immunorecognition proteins etc). Despite this flexibility, however, geometric validation in COOT (Emsley *et al.*, 2010) indicated that all torsion angles remain within acceptable parameters and it is therefore unlikely that LdcB exploits ligand distortion to aid in catalysis. One notable exception was identified in the PDB file of a peptidoglycan recognition protein (2CB3), which contained a ligand that was constructed as a single, unique molecule that required a custom restraint file, rather than being constructed of defined sugars and amino acids (MurNAc, L-Ala, D- γ -Gln, L-

Lys [or meso-DAP], D-Ala). Manual inspection of the torsion angles between each amino acid in COOT revealed that it, like the other ligands, has been constructed with the correct geometries.

The major difference between many of the peptidoglycan ligands found in the literature and the M-tri-Lys(D-Asn) ligand that co-crystallised with *SpLdcB* is in the conformation of the lysine component, as mentioned previously, which is found wrapped around the active site rather than in the more common extended conformation. It is, however, unlikely that this conformation is true to the biological role of the enzyme, as expanded upon above. Interestingly, many of the ligands were observed to superpose onto one another identically at the MurNAc sugar and N-terminal L-alanine residues, with flexibility being imparted by the relatively long D-γ-glutamine residue, and in part by the lysine or meso-DAP sidechain.

Several key structural changes occur upon ligand binding in *SpLdcB* (**Figure 4.15**, overleaf). First and foremost, substrates are occluded in the empty form of the enzyme by a loop between residues Gly163 and Glu171, which blocks access to the active site. Most of these residues (Asp167 to Glu171) move by 10-15 Å in the presence of ligand to open up the cleft by re-folding into an additional turn at the N-terminus of α-helix 4. The re-arrangement of these residues is accompanied by further conformational changes in Trp206, the indole ring of which moves by up to 9 Å, which not only aids to open up the active site of the enzyme, but also allows the side chain of Glu204 to re-arrange to take up a position capable of interacting with the substrate. It is unknown if this movement occurs prior to or post-catalysis. The importance of Glu204 and the zinc ion for catalysis is highlighted by the observation that tetra-D is a much poorer substrate for the E204A variant of *SpLdcB*, and the absence of activity in the presence of EDTA (**Section 4.3**). The conformational changes seen on ligand binding to *SpLdcB* contrasts with the absence of significant conformational changes described for VanXYg in complex with phosphinate (PDBid: 4muq (Meziane-Cherif *et al.*, 2014)). However, it should be noted that the 'closed' confirmation of *SpLdcB* also has a small molecule bound in the active site, and VanXYg may undergo conformational changes when interacting with a larger, pentapeptide substrate.

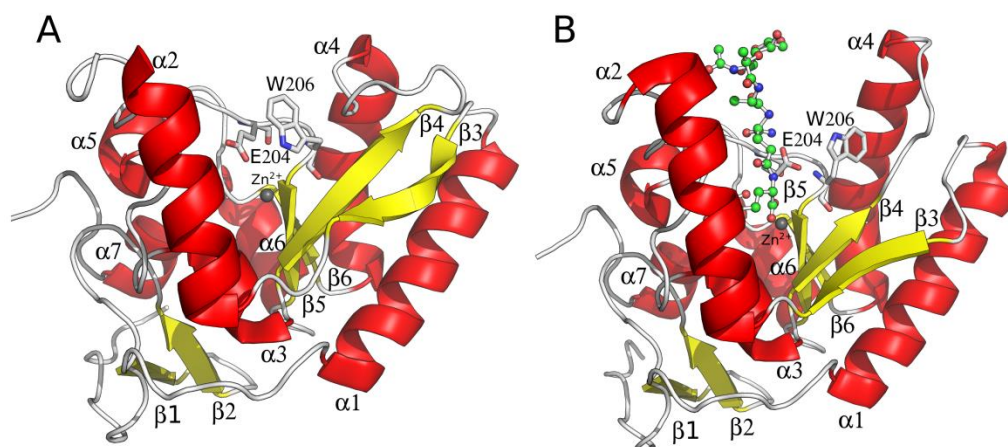


Figure 4.15: Conformational changes in *SpLdcB* on ligand binding. (A, B) Cartoon representations of apo *SpLdcB* (A) and *SpLdcB* with bound MurNAc-Ala-D-γ-Gln-Lys-(D-Asn) (B); α-helices (red) and β-strands (yellow) are numbered. The loops between secondary structure elements are coloured silver. The bound zinc ion is shown as a grey sphere. Key residues that move upon ligand binding are shown in stick representation.

B-factor analysis of *SpLdcB* in both the opened and closed conformations reveals that the mobile, substrate binding loop has the highest thermal motion in the open conformation (**Figure 4.16**), perhaps indicating why many of the LdcB structures solved are in the closed conformation. It should be noted that the portion of the binding loop that folds into the N-terminus of α-helix 4 is more stable in the open conformation, as would be expected due to hydrogen bonds formed with the remainder of the helix. In addition, a portion of the C-terminus of α-helix 2 could not be built in the open conformation of LdcB, suggesting disorder which is confirmed by B-factor analysis. The entire α-helix was modelled in the closed conformation of *SpLdcB*, suggesting greater order. It is unclear as to whether the increased stability of α-helix 2 is due to ligand binding or is the result of a change in crystal contacts between native and liganded *SpLdcB*.

4.2.4 Modelling of the tetrapeptide substrate

Using the structure of *SpLdcB* with M-Tri-Lys(D-Asn) bound, together with the structure of *SpLdcB* that was solved with D-Ala bound in the active site and the structures of VanXYc and VanXYg with D-Ala bound, a model of *SpLdcB* with the entire tetrapeptide ligand bound can be proposed (**Figure 4.17**, overleaf). In this model, the contacts made by the MurNAc sugar, L-Ala, D-γ-Gln and L-Lys are maintained, with the side-chain of L-Lys occupying the path proposed in **Section 4.2.3**. The D-alanine has

been modelled in at the S1' subsite, replacing the D-Asn of the M-Tri-Lys(D-Asn) product mimic and superposing well onto the D-alanine residue found to co-purify with *SpLdcB*.

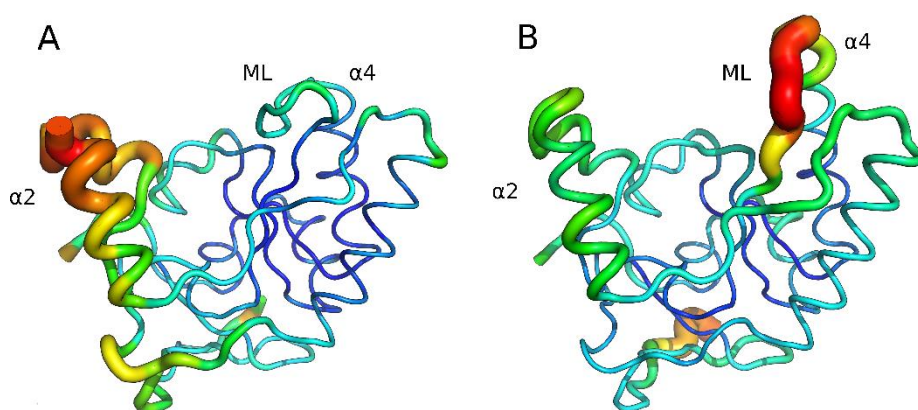


Figure 4.16: B-factor analysis of *SpLdcB* in the closed (A) and open (B) conformations. B-factor values are shown both by thickness of the tube, with thicker sections having higher B-factors, and by the colour of the tube; areas of high B-factors are shown in red, with lowest B-factors shown in blue. α -helix 2 and 4 are highlighted, as it the mobile loop that changes conformation between forms, as these areas show the greatest change in B-factor values between the two conformations.

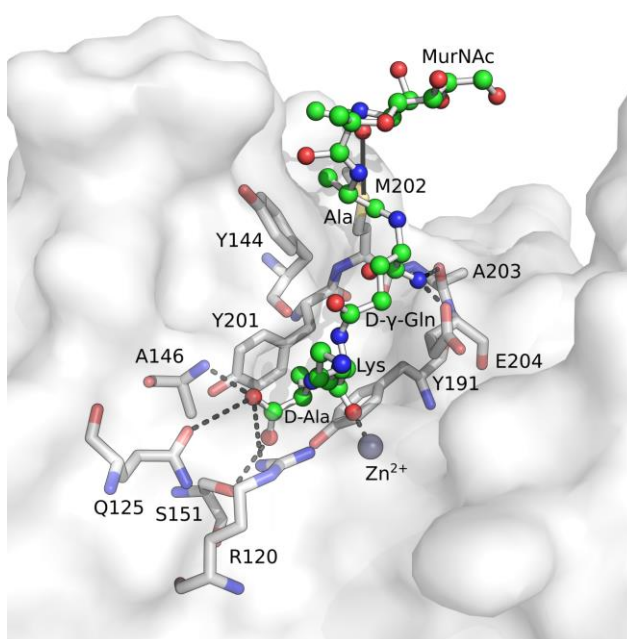


Figure 4.17: The modelled path of the tetrapeptide substrate of *SpLdcB*. Met202 interacts with the MurNAc sugar N-acetyl group via Van der Waals forces whilst Tyr144 stacks against the ligand Ala to sterically occlude D-Ala at this point. Ala203 and Glu204 interact with D- γ -Gln. The Lys residue of the ligand coordinates to the zinc ion, completing the zinc's tetrahedral coordination sphere, with the side-chain of the Lys projecting out of the active site. Tyr201 and Tyr191 stack against the terminal D-Ala, which is in turn hydrogen bonded to Ala146, Gln125, Ser151 and Arg120. The tetrapeptide ligand (MurNAc-Ala-D- γ -Gln-Lys-D-Ala) is shown as ball-and-stick representation, the key interacting residues from *SpLdcB* as sticks with white carbons and the surface of *SpLdcB* is shown in transparent white. The zinc ion is shown as a grey sphere. Hydrogen bonds made by the tetrapeptide ligand with *SpLdcB* are shown as black dashed lines, whilst van der Waals forces are shown as a solid silver line.

4.3 Biochemical and biophysical assays on the native enzyme

4.3.1 Thin-layer chromatography (TLC) of *Sp*LdcB reaction products

To confirm that LdcB enzymes are LD-carboxypeptidases, the enzymes were tested for activity against the synthetic tetrapeptides tetra-D (L-Ala—D-Gln—L-Lys—D-Ala) and tetra-L (L-Ala—D-Gln—L-Lys—L-Ala). The two peptides differ from the natural cell wall tetrapeptide in two ways; first, they lack the N-terminal MurNAc sugar moiety and second, D-γ-Gln is replaced by D-Gln (**Figure 4.18**).

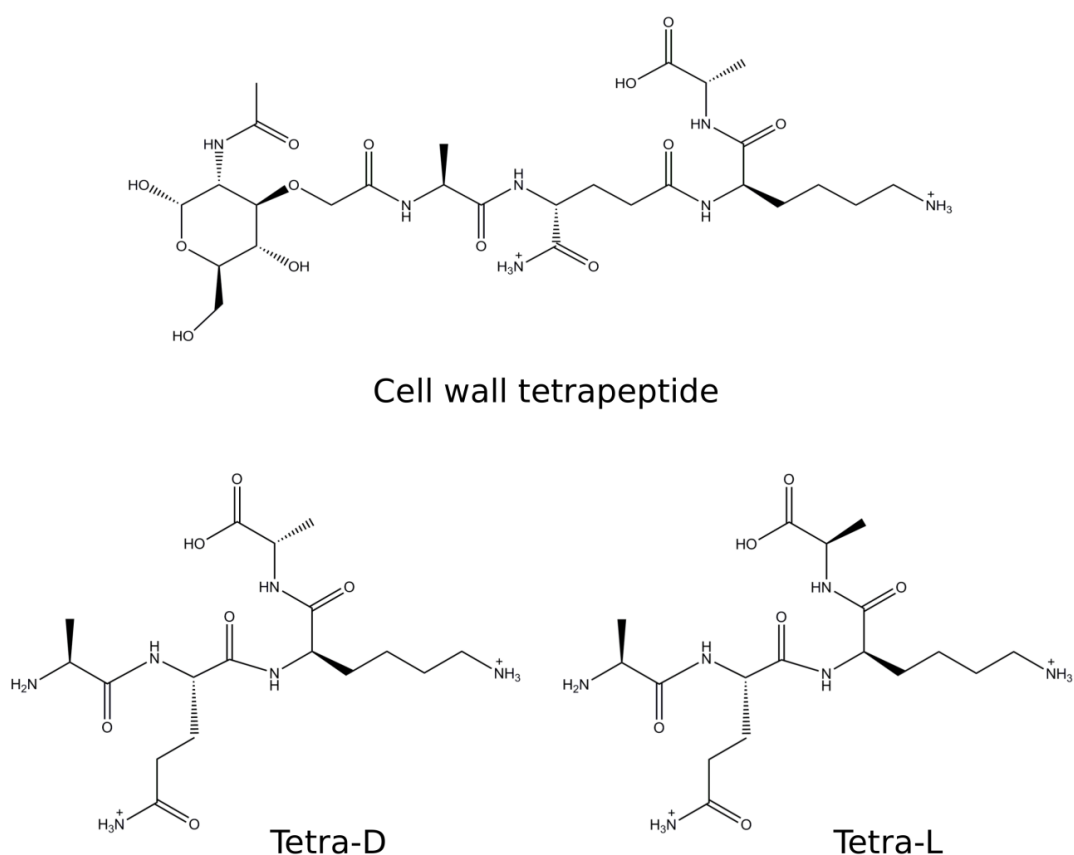


Figure 4.18: Chemical structures of substrates used in this study. Tetrapeptide substrates used for catalytic studies (Tetra-D, Tetra-L), compared to the true substrate (cell wall tetrapeptide).

Whilst the changes from the natural substrates most likely results in lower binding affinities for the enzyme, the two peptides do allow for discrimination between LD- and LL-carboxypeptidase activities. Following a 2 hour incubation of the protein

with the peptides, in the presence of either zinc or EDTA, the reactions were spotted onto a silica TLC plate to allow the products to be resolved. After optimum separation had been achieved, the TLC was stained with ninhydrin, which reacts with terminal amines and lysine sidechains to stain the products. The TLCs revealed that *SpLdcB* was active against only the tetra-D ligand, and that this activity was dependent on the presence of zinc; activity was abrogated in the presence of EDTA (**Figure 4.19**, overleaf).

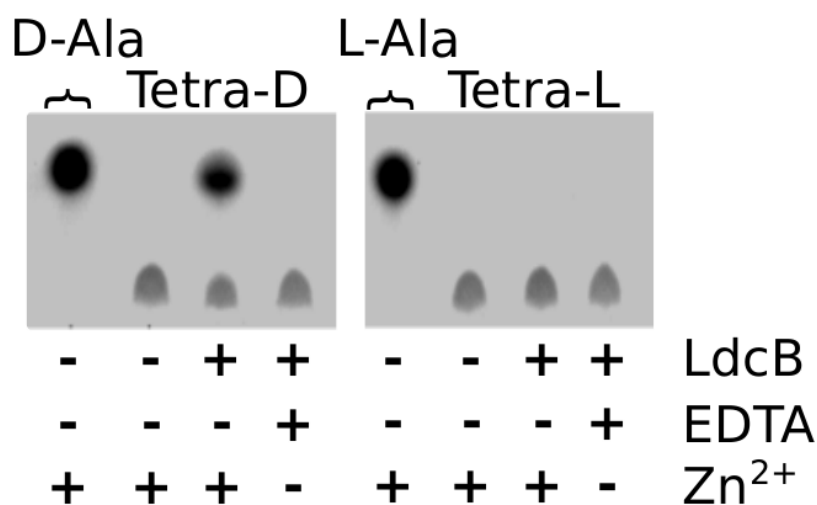


Figure 4.19: Activity of *SpLdcB* against peptidoglycan tetrapeptides. Thin layer chromatography (TLC) of *SpLdcB* incubated with various substrates. D-alanine is less polar than the tetrapeptide and migrates faster in TLC. *SpLdcB* is able to hydrolyse only the Tetra-D substrate, and only in the presence of zinc.

These results agreed with those obtained by Dr Christine Aldridge, Newcastle University (Hoyland *et al.*, 2014), in which *SpLdcB* was incubated against cell wall material isolated from an *ldcB* deficient strain of *S. pneumoniae* (*S. pneumoniae* R6Δ*ldcB*). The *ldcB* mutant exhibited an increased amount of tetrapeptides present within the cell wall, together with a co-committal decrease in tripeptide content; incubation with *SpLdcB* reversed this affect, but not in the presence of EDTA (**Figure 4.20**, overleaf) (Hoyland *et al.*, 2014).

In addition to cell wall tetrapeptides, LdcB enzymes were tested for activity against D-ala-D-ala dipeptide and cell wall pentapeptide (Ala-D-γ-Gln-Lys-D-Ala-D-Ala). *SpLdcB*, *BsLdcB* and *BaLdcB* (plasmid provided by Dr. Djalal Meziane-Cherif, Institut Pasteur,

France) were incubated against the three substrates in the presence of zinc and the reaction products separated by TLC. All three enzymes showed activity against the D-Ala-D-Ala dipeptide and the tetrapeptide, while only *SpLdcB* showed weak activity against the pentapeptide (**Figure 4.21**, overleaf).

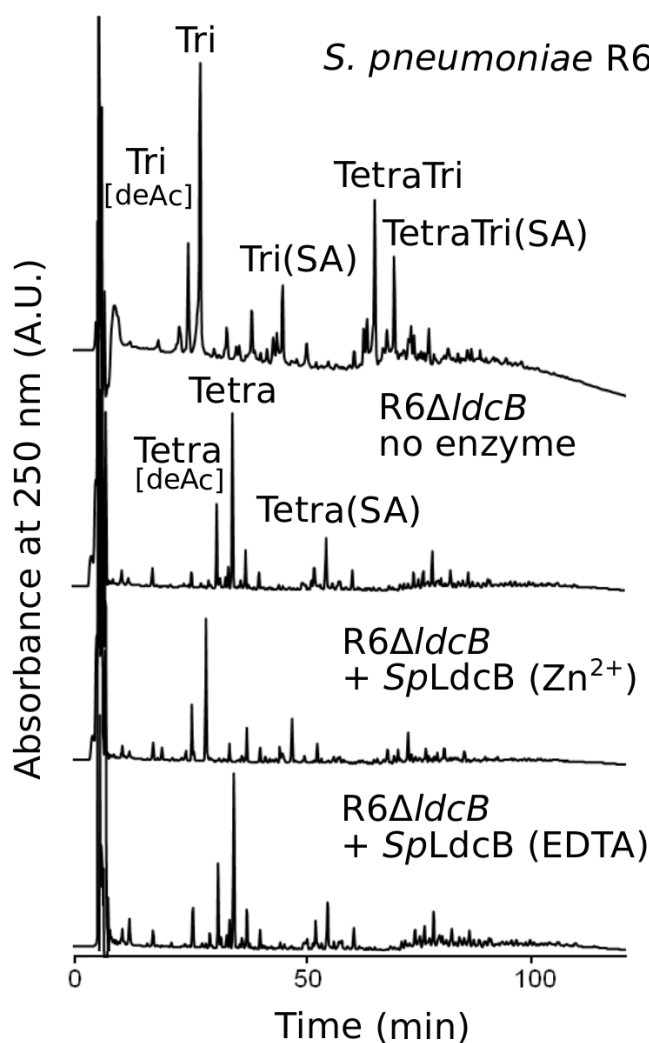


Figure 4.20: HPLC chromatograms of mucopeptides from *S. pneumoniae* strains, obtained after incubating PG with or without *SpLdcB* in the presence of Zn^{2+} or EDTA, followed by digestion with cellosyl and reduction with sodium borohydride. Figure taken from Hoyland *et al.*, (2014), performed by Dr Christine Aldridge, Newcastle University. The mucopeptides are: Tri, GlcNAc—MurNAc(r)—L-Ala—D-γ-Gln—L-Lys; Tetra, GlcNAc—MurNAc(r)—L-Ala—D-γ-Gln—L-Lys—D-Ala; TetraTri, GlcNAc—MurNAc—L-Ala—D-γ-Gln—L-Lys—D-Ala—L-Lys—D-γ-Gln—L-Ala—MurNAc(r)—GlcNAc; TetraTetra, GlcNAc—MurNAc—L-Ala—D-γ-Gln—L-Lys—(D-Ala)—D-Ala—L-Lys—D-γ-Gln—L-Ala—MurNAc(r)—GlcNAc; deAc, deacetylation at GlcNAc; SA, L-Ser—L-Ala branch (at L-Lys). GlcNAc, *N*-acetylglucosamine; MurNAc(r), *N*-acetylmuramic acid.

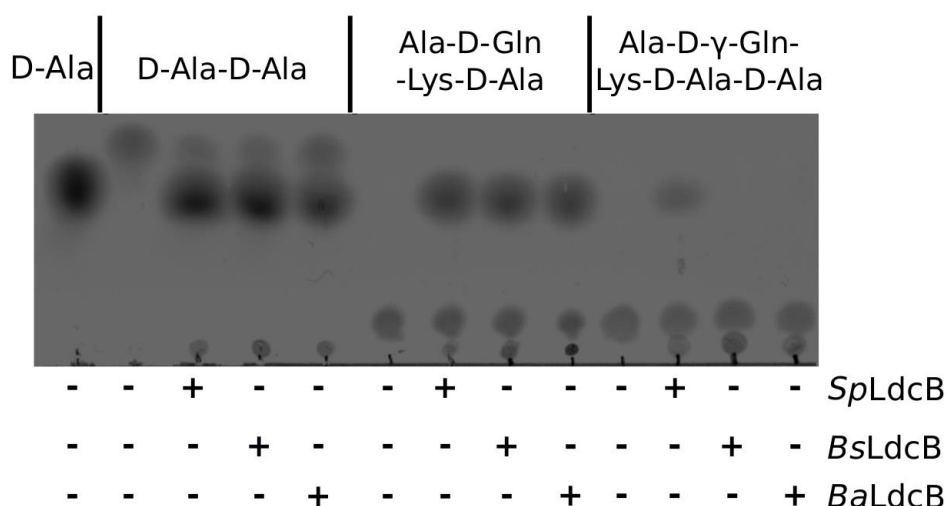


Figure 4.21: Activity of LdcB homologues against peptidoglycan pentapeptide and D-Ala-D-Ala dipeptide. Thin layer chromatography (TLC) of *SpLdcB*, *BsLdcB* and *BaLdcB* incubated with various substrates. D-alanine is less polar than the tetrapeptide and migrates faster in TLC. All three homologues are able to hydrolyse the tetrapeptide and D-Ala-D-Ala dipeptide substrates, but only *SpLdcB* is active against the pentapeptide.

Although the ability of LdcB to hydrolyse the D-Ala-D-Ala dipeptide was unexpected, it is not surprising that the small substrate could enter an active site designed to accommodate tetrapeptides. It is likely that the long incubation time of the reaction, in addition to the high substrate concentration used, resulted in the hydrolysis of the dipeptide. Likewise, it is likely that the hydrolysis of the pentapeptide by *SpLdcB* is a reaction artefact; it has been shown that deletion of the DD-pentapeptidase *dacA*, responsible for hydrolysing cell wall pentapeptides to tetrapeptides in *S. pneumoniae*, causes the accumulation of pentapeptides (Barendt *et al.*, 2011) that would not occur if LdcB was highly active against this substrate.

Modelling the pentapeptide substrate in the active site of *SpLdcB* revealed potential interactions that might be necessary for catalysis. In the model, many of the observed interactions between the ligand and the protein were lost, possibly explaining the much weaker activity. Several observed contacts were maintained or new ones made, however (**Figure 4.22**, overleaf). The terminal D-Ala retained many of the interactions that were made by the terminal D-Ala of the tetrapeptide, losing only the hydrogen bond to Ala146. The carbonyl oxygen of the penultimate D-Ala completed the coordination geometry of the zinc ion in place of the L-Lys, whilst the L-Lys was moved upwards one position, with the carbonyl oxygen now placed within hydrogen bonding distance of Glu204. Slight rearrangement of the sidechain of Asp136 could place it

within hydrogen bonding distance of the N_ϵ of the lysine sidechain, though this interaction could not be made with cross-linked material. Finally, the sidechain of D- γ -Gln was able to interact with Glu205, with neither MurNAc nor L-Ala contacting the protein. In VanXYc, the O1 of the MurNAc sugar would be positioned to interact with Arg119, stabilising the ligand further and possibly explaining the preference of this enzyme for pentapeptides; this prediction is based upon the closed conformation of the bisubstrate selectivity loop, and thus may not be valid for VanXYc complexed with the pentapeptide.

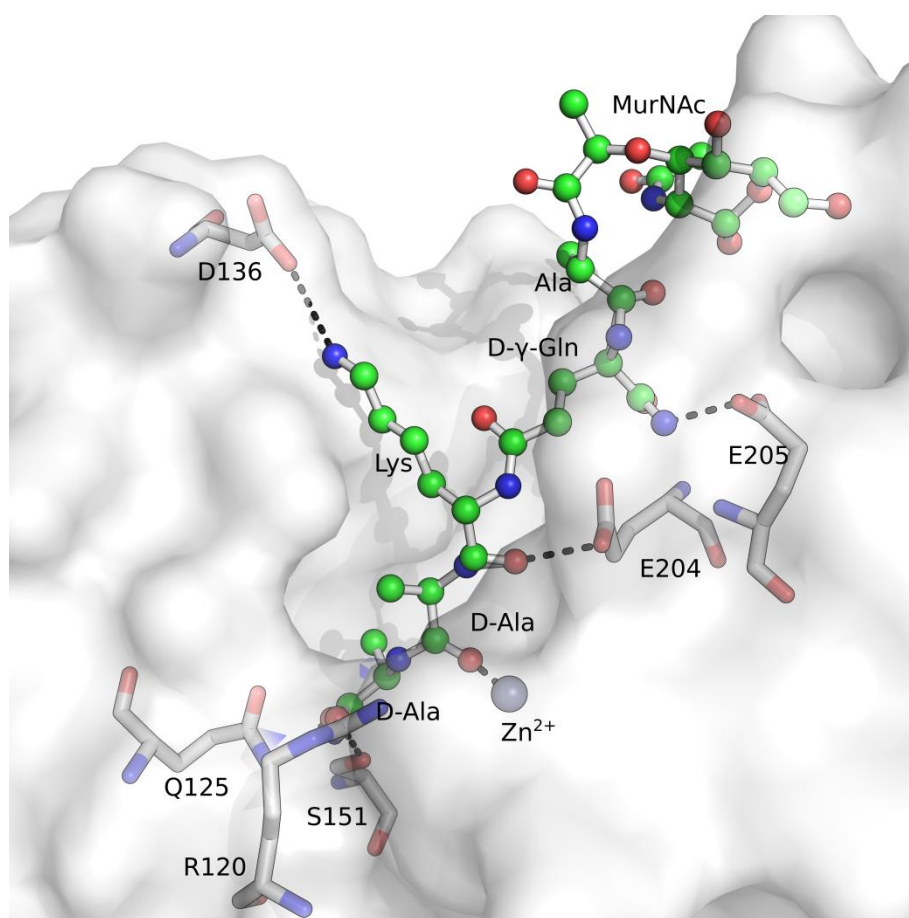


Figure 4.22: The predicted interactions of MurNAc-pentapeptide with *SpLdcB*. The pentapeptide was based on the structure of *SpLdcB* with M-Tri-Lys(D-Asn) bound. The tetrapeptide ligand was moved as a rigid body and two D-Ala residues added to the C-terminus. The D-Asn moiety was not present on the pentapeptide hydrolysed by *SpLdcB* in TLC experiments and so was removed from the model. The ligand is shown in ball and stick representation, with interacting amino acids shown as stick with white carbons. The overall structure of the protein is shown as a surface, whilst the zinc ion is shown as a grey sphere. Potential hydrogen bonds made by the pentapeptide ligand with *SpLdcB* are shown as black dashed lines.

In both *BsLdcB* and *BaLdcB*, two key interactions predicted by the *SpLdcB*-pentapeptide model were not possible; Glu205 was replaced by proline whilst lysine was present in place of Asp136. The inability of these amino acids to stabilise the protein-ligand complex likely could result in weaker binding of the pentapeptide to perhaps explain the lack of activity of *BsLdcB* and *BaLdcB* against the pentapeptide.

4.3.2 Construction and characterisation of active site mutants

The crystal structure of *SpLdcB* with bound ligand allowed for the identification of two residues, Arg120 and Glu204, that might participate in substrate recognition and/or catalysis. Both Glu204 and Arg120 are strictly conserved in LdcB sequences, in addition to VanY, VanXYg and VanXYc. Indeed, both residues are found in the active sites of several zinc-carboxypeptidases (Christianson and Lipscomb, 1989; Odintsov *et al.*, 2004). Glu204 is positioned to activate the zinc-bound water for nucleophilic attack, whilst Arg120 is capable of contributing to the polarisation of the carbonyl oxygen of the tetrapeptide, preparing it for nucleophilic attack.

To confirm the importance of these residues in LdcB enzymes, Quikchange site-directed mutagenesis was used to produce alanine substitutions. Additionally, an E204H mutant was constructed in an effort to probe the differences between E-x-x-H and H-x-H members of the LAS family of enzymes.

4.3.2.1 Confirmation of mutant folding

Prior to embarking on activity assays utilising the mutant forms of *SpLdcB*, circular dichroism (CD) was used to confirm that the mutageneses had no impact on protein folding. A CD scan from 250 - 190 nm revealed that all mutant proteins were folded similar to wild-type *SpLdcB* (**Figure 4.23**), with minor differences below 200 nm likely due to an increase in noise in this region (Kelly and Price, 2000). The mutated proteins were thus entirely suitable for activity assays.

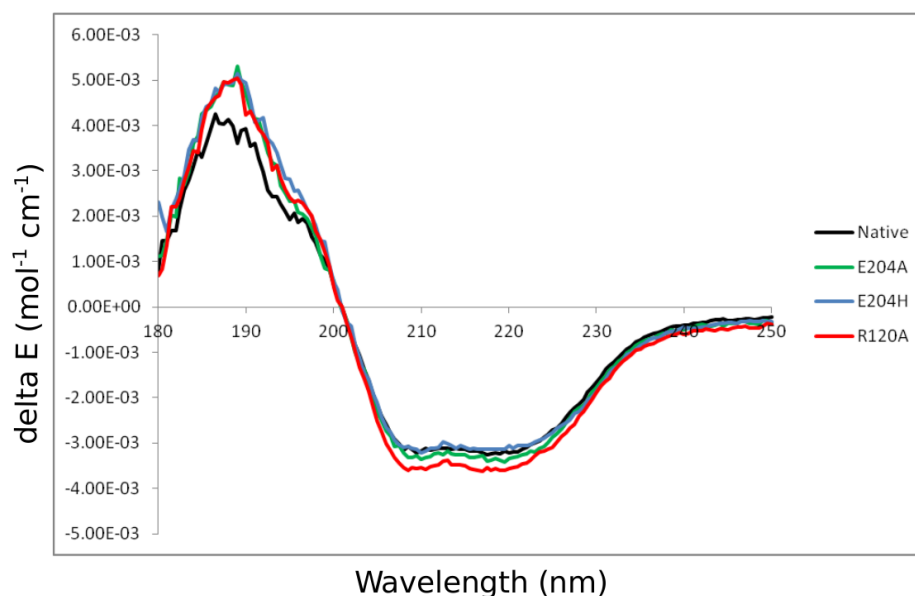


Figure 4.23: Circular dichroism (CD) spectra of *SpLdcB* and three mutants. The overall shape of each spectrum is identical, with mostly equal peak intensities, indicating that the three mutants fold in a similar manner to that of the wild-type.

4.3.2.2 Thin-layer chromatography (TLC) of mutant *SpLdcB* reaction products

To assess the impact the single mutations had on *SpLdcB*, the mutants (700 μ M) were each incubated against the tetra-D substrate (20 mM) for 2 hours in the presence of a molar equivalent of zinc. The reactions were spotted onto a silica TLC plate, the products resolved and the TLC stained with ninhydrin. Only the *SpLdcB* variants E204H and E204A appeared to have any noticeable effect on catalysis (**Figure 4.24**).

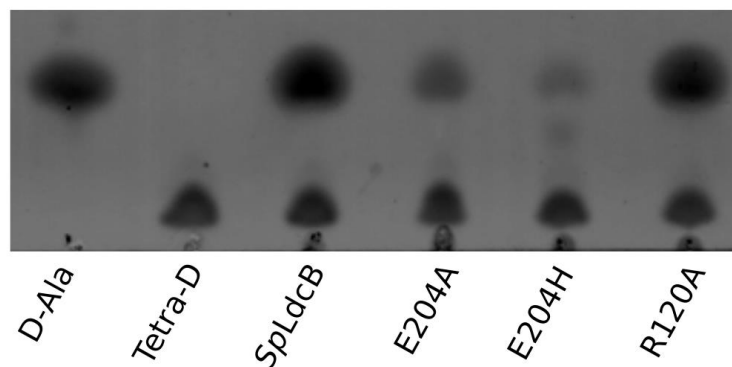


Figure 4.24: Activity of wild-type (WT) *SpLdcB* compared to three active site mutants. Thin layer chromatography (TLC) of WT and mutant *SpLdcB* incubated with Tetra-D. *SpLdcB* activity appears unaffected by the R120A mutation, while the E204A and E204H mutants experience a severe decrease in activity.

To resolve the differences between the E204H and E204A mutants and wild-type *SpLdcB* a time-course was performed using lower concentrations of protein (50 μ M), from which samples were taken at 15 minute intervals and the reaction stopped by the addition of ethanol to a final concentration of 50%. The reaction products were resolved by TLC as described above, with both the E204A and E204H showing reduced catalytic activity (**Figure 4.25**, overleaf). The E204A mutant showed slightly lower catalytic activity compared to the wild-type protein, producing similar levels of D-Ala only after 30 minutes, while the E204H mutant was significantly slower.

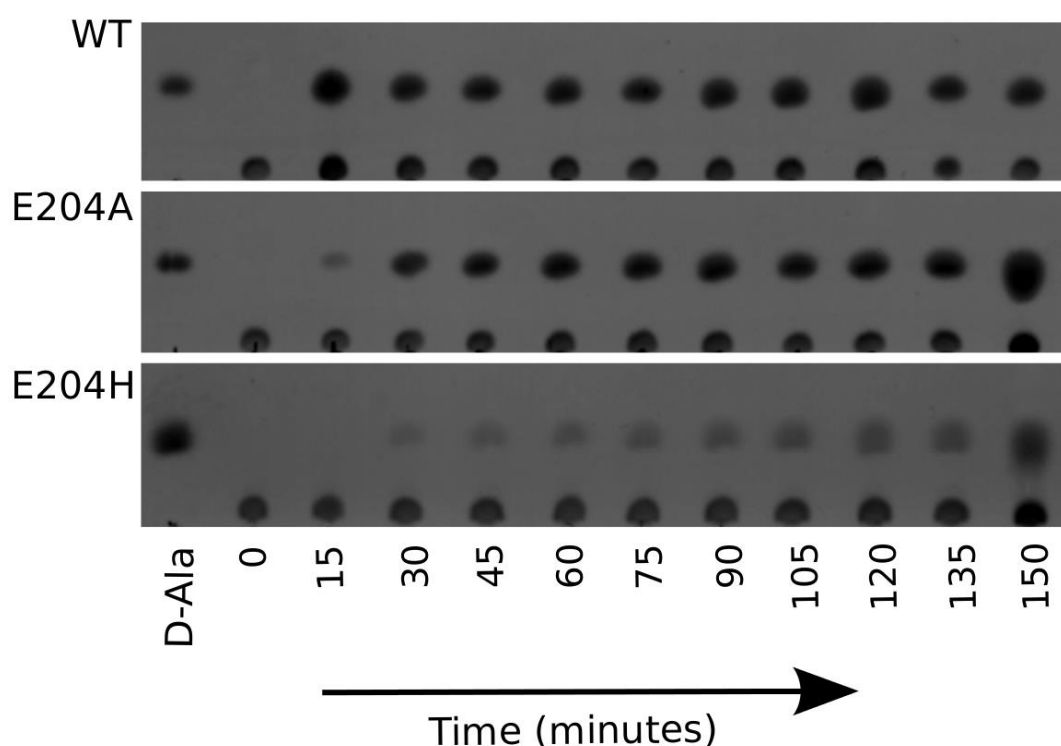


Figure 4.25: Time-course TLC of WT, E204A and E204H *SpLdcB*. Thin-layer chromatography (TLC) of WT, E204A and E204H *SpLdcB* incubated with Tetra-D. Samples were taken at 15 minute intervals and the reaction halted with the addition of ethanol. *SpLdcB* activity was most affected by the E204H mutant.

To relate the activity of the active site mutants to the structure of *SpLdcB*, *in silico* mutagenesis was performed in COOT (Emsley *et al.*, 2010) on both the apo and ligand-bound *SpLdcB* structures (**Figure 4.26**, overleaf). The mutation of Glu204 to alanine resulted in a sidechain that could neither participate in catalysis nor form hydrogen bonds with the D- γ -Gln moiety of the substrate. The reduced catalytic rate likely results from effects on both K_m and V_{max} . Like the E204A mutant, the E204H

mutant appeared to be unable to participate in catalysis, as all the conformations that positioned the histidine near the zinc ion resulted in steric clashes with other active site amino acids. Indeed, the only conformation that the histidine could occupy was near the path of the incoming ligand (within 1.4 Å of the sidechain), likely causing conformational changes in the ligand and/or the protein, reducing the K_m .

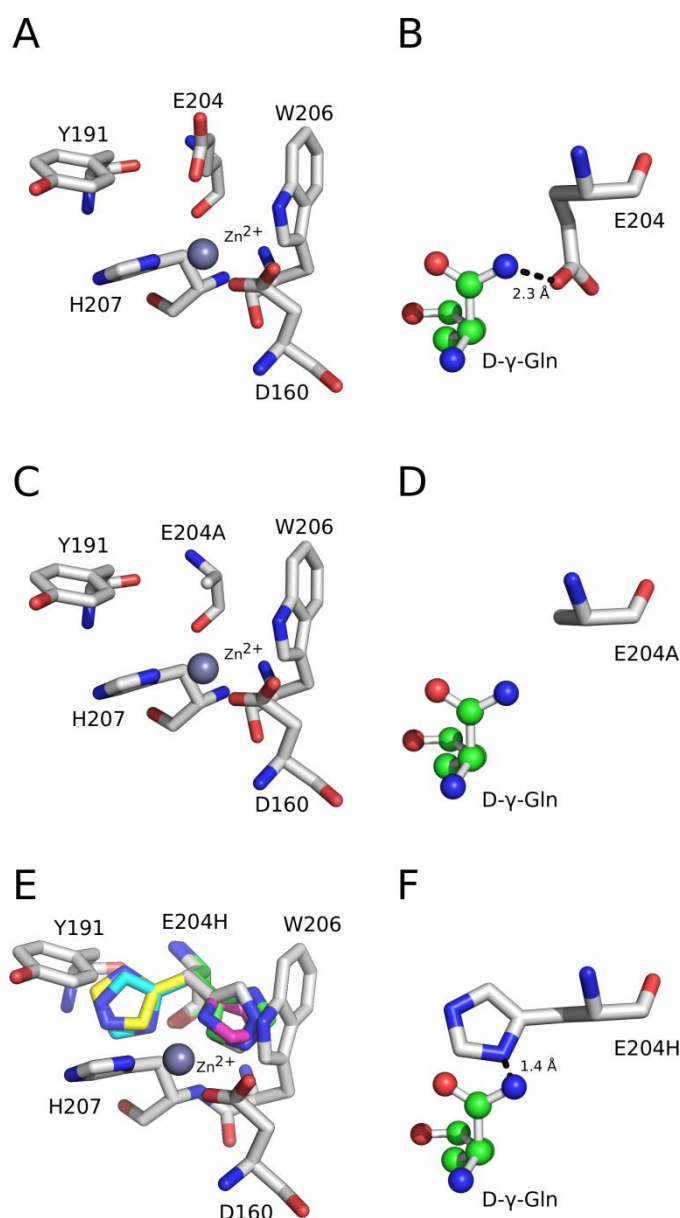


Figure 4.26: *In silico* mutagenesis of *SpLdcB* E204A and E204H. Panels A and B show the native enzyme; no clashes between E204 and the mainchain of *SpLdcB* or the M-Tri-Lys(D-Asn) ligand can be seen. Panels C and D show the E204A mutant; the alanine does not clash with any other amino acids, mainchain or ligand, but cannot interact with the ligand. Panels E and F show the E204H mutant; the histidine clashes with either the main protein chain (E) or the ligand (F) in many sidechain orientations. Histidine 102 is not shown coordinating the zinc ion in panels A, C & E, and only D-γ-Gln of the ligand is shown in panels B, D and F for clarity. The ligand is shown in ball and stick representation, the zinc ion as a grey sphere and protein amino acids as sticks with white carbons, with the exception of E204H in panel E, in which the carbons have been coloured differently to allow the various rotomers to be distinguished.

4.4 Kinetic characterisation of *SpLdcB* for Ala-D-γ-Gln-Lys-D-Ala

Though the TLC experiments confirmed that LdcB enzymes have LD-carboxypeptidase activity, a quantitative assay was required for a more detailed analysis. Whilst the release of D-alanine is not measurable directly, linked assays have been used previously to monitor peptidase reactions (McKevitt *et al.*, 2007).

SpLdcB, *BsLdcB* and *BaLdcB* (overexpression plasmid provided by Dr. Djalal Meziane-Cherif, Institut Pasteur, France) were purified, the native metal removed by dialysis against 50 mM EDTA, and reloaded with zinc at a 1:1 stoichiometry following removal of the EDTA. Proteins were flash-frozen in liquid nitrogen and provided to collaborators in the Institut Pasteur (Paris, France) for kinetic analysis. All kinetic experiments were performed by Dr. Djalal Meziane-Cherif and were based upon the amino-acid oxidase-lactate dehydrogenase coupled assay (Araoz *et al.*, 2000). In brief, the D-alanine liberated from the tetrapeptide by the action of an LD-carboxypeptidase is converted into pyruvate by D-amino acid oxidase, with pyruvate then reduced by lactate dehydrogenase (LDH) in a NADH dependent manner. The conversion of NADH to NAD⁺ was followed at 340 nm in an Uvikon UV931 spectrophotometer to monitor the velocity of the LD-carboxypeptidase (**Figure 4.27**, overleaf). Steady-state kinetic parameters were obtained by fitting the data to the Michaelis-Menten equation in the program Sigmaplot v11.0 (Systat Software Inc. Unfortunately, due to the limited availability of the substrate (Ala-D-γ-Gln-Lys-D-Ala), a custom synthesis by Activotec (Cambridge, UK), saturation could not be achieved for *SpLdcB* or *BaLdcB*, and only one data point for *BsLdcB* at saturation was obtained. Additionally, only two repeats could be performed for each enzyme. Thus, the kinetic parameters determined by this assay are only estimates and more work is required to establish their values more accurately.

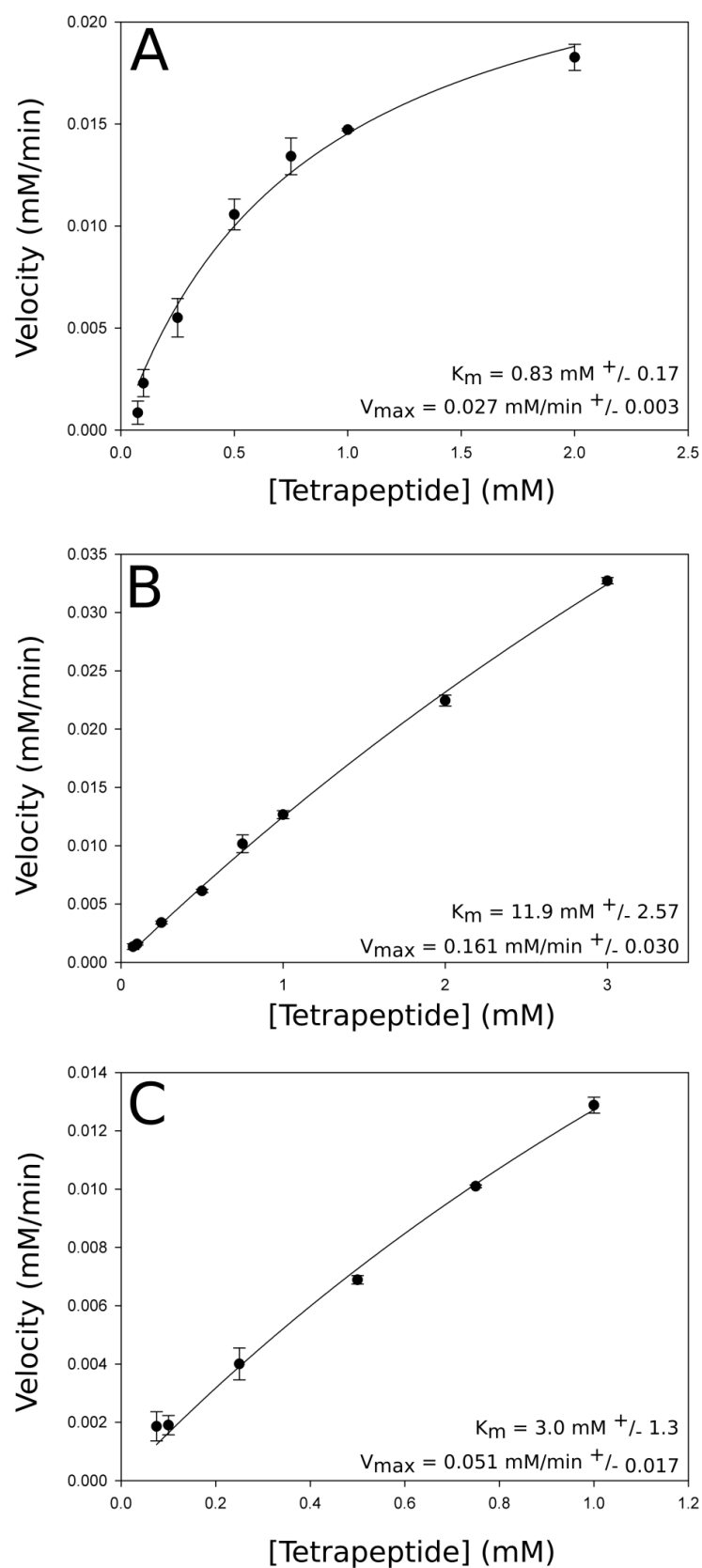


Figure 4.27: Michaelis-Menten curves for *BslDcB* (A), *SplDcB* (B) and *BalDcB* (C). All curves were plotted in Sigmaplot v11.0 from data provided by Dr. Djalal Meziane-Cherif, Institut Pasteur, France.

The large difference in substrate affinity and turnover rates between the three LdcB enzymes is unlikely to be explained by their substrate specificities. The tetrapeptide used in the assay is identical to that found in the cell wall of *S. pneumoniae*, whereas the cell walls of the Bacilli contain D-γ-Glu in place of D-γ-Gln position 2 and meso-Diaminopimelic acid (meso-DAP) in place of L-Lys at position 3, therefore, *SpLdcB* would be expected to exhibit the strongest K_m . If the differences between the enzymes' K_m and k_{cat} values are not related to the experimental constraints, the differences might be related to the enzymes' cellular functions. *SpLdcB* appears to be required for division in *S. pneumoniae* (Barendt *et al.*, 2011), and an overactive enzyme is likely to result in aberrant positioning of division septa. *BsLdcB* appears to be used for D-alanine recycling and growth under D-alanine limited conditions (Hoyland *et al.*, 2014). Under such conditions, the tight affinity of *BsLdcB* for the tetrapeptide could permit efficient cleavage to result in a large pool of D-alanine liberated for peptidoglycan synthesis.

The K_m obtained for all three LdcB enzymes, though only estimates, are comparable to other carboxypeptidases, including VanXY_c for UDP-MurNAc-pentapeptide (1.39 mM \pm 0.12; (Meziane-Cherif *et al.*, 2014) and carboxypeptidase Y (0.07 – 16.5 mM, depending on the substrate and pH; (Hayashi *et al.*, 1975).

4.5 A Proposed catalytic mechanism

The catalytic mechanism of zinc metallo-carboxypeptidases has been the subject of several decades-worth of mechanistic studies, with work focusing on the archetypical enzyme carboxypeptidase A (CPA) (Koshland, 1958; Folk and Gladner, 1960; Christianson and Lipscomb, 1989; Wu *et al.*, 2010). The mechanism proposed for the hydrolysis of C-terminal peptide bonds by CPA is known as the promoted-water pathway (Wu *et al.*, 2010) and involves Glu270 acting first as a general base to activate a zinc-bound water. The activated water attacks the scissile carbonyl carbon, which has been polarised by the effects of the zinc ion and Arg127, before acting as a general acid to transfer a proton to the leaving amino group (Wu *et al.*, 2010).

Although CPA is structurally distinct from *SpLdcB* (RMSD of 5.68 Å over 100 aligned Cαs), their active sites share remarkable similarity, essentially being mirror images of each other (**Figure 4.28**). Both enzymes contain a zinc ion co-ordinated by two

histidines (His69, His196 in CPA; His153, His207 in *SpLdcB*) and an acidic amino acid (Glu72 in CPA; Asp160 in *SpLdcB*) opposite an arginine (Arg127 in CPA; Arg120 in *SpLdcB*). In both proteins the arginine and glutamate/aspartate straddle a water molecule, but their positions relative to their counterparts are reversed (**Figure 4.26**). Consequently, the path of the substrate through the active site is also reversed. Given the retention of the metal co-ordination and key adjacent amino acids in LdcB, it is entirely plausible that the reaction mechanism of LdcB occurs in a similar fashion to that of CPA.

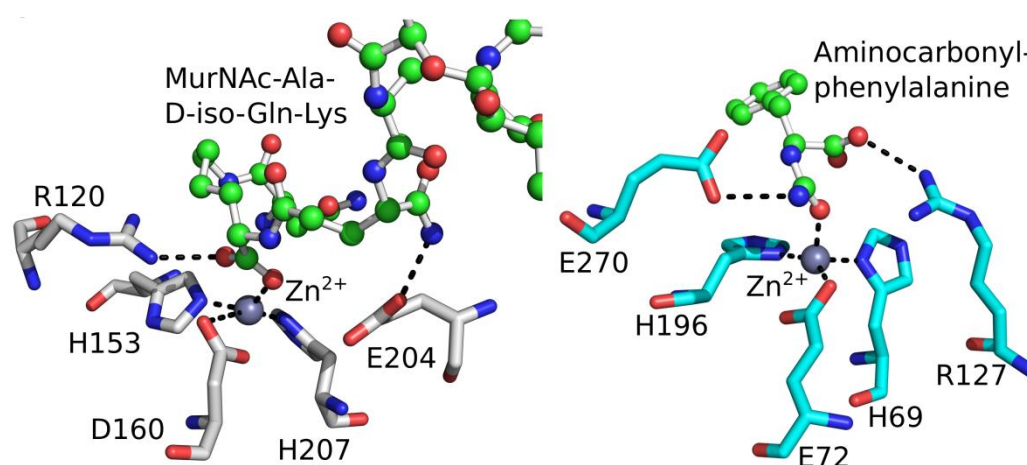


Figure 4.28: The active sites of *SpLdcB* and pancreatic carboxypeptidase A (PDB ID: 1HDU (Cho *et al.*, 2002) are compared. In each structure, the Zn^{2+} ion is co-ordinated by two histidines (His69, His196 in carboxypeptidase A; His153, His207 in *SpLdcB*) and an acidic amino acid (Glu72 in pancreatic carboxypeptidase A; Asp160 in *SpLdcB*). The reversed locations of Arg127 and Glu270 in pancreatic carboxypeptidase A relative to Arg120 and Glu204 in *SpLdcB* is matched by a reversal of the path of the respective substrates through the active sites, and the likely retention of a common catalytic mechanism.

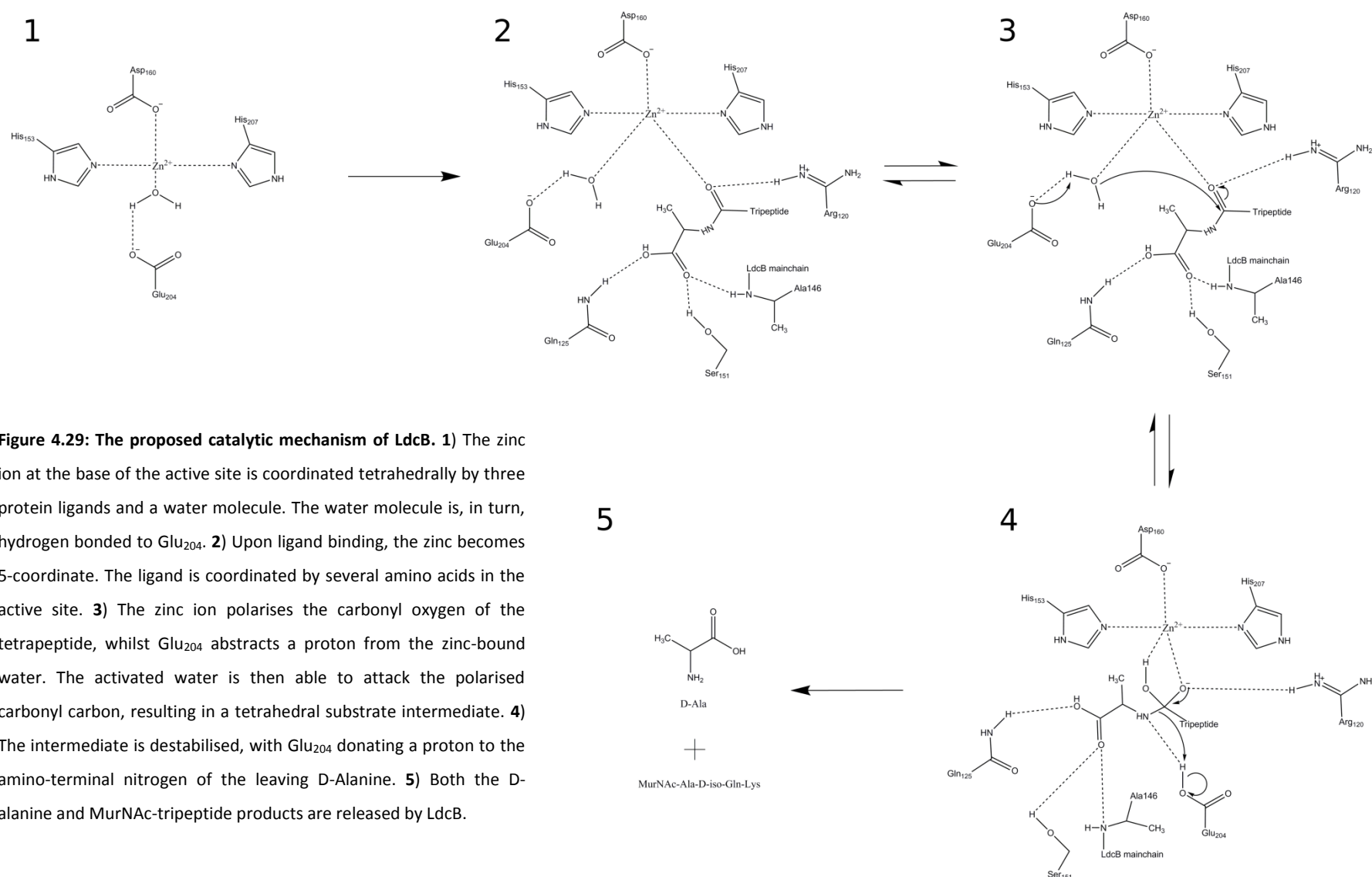
The use of the promoted-water pathway by LdcB is supported by the effect the E204A mutant has on catalysis; mutating glutamate to alanine results in an amino acid with a shorter side-chain that is not ionisable and is non-polar, preventing the amino acid acting as a general base to activate the catalytic water (**Figures 4.25 & 4.26**). The lowering of the catalytic rate, rather than abolishment of catalysis, in the E204A mutant likely arises from the zinc ion acting as the general base to activate the catalytic water, as has been previously postulated for CPA (Christianson and Lipscomb, 1989; Kilshtain-Vardi *et al.*, 2002). In addition, the E204A mutant is likely to possess a different pH profile compared to the wild-type enzyme. As with several other zinc

metalloproteins that function via a catalytic glutamate, such as CPA (Mock and Tsay, 1988) or leucine aminopeptidase (Bzymek and Holz, 2004), LdcB is likely most active in a pH range of ~5 to 10, with activity rapidly decreasing as the pH approaches and is reduced below 4.07, the pKa of glutamate, due to protonation of the catalytic glutamate. It should be noted however, that the presence of the catalytic zinc may play a role in altering the local environment of the glutamate and so the enzyme may be active over a larger pH range than would be expected from an active site comprised of just the glutamate. Studies of enzymes that utilise a zinc ion for catalysis, such as carbonic anhydrase (Lindskog and Coleman, 1973), have shown that these enzymes have optimal activity around pH 7, with activity reducing at lower pH values. Therefore, the E204A mutant, in which the zinc acts as the catalytic component of the enzyme, would likely have a much narrower activity range with respect to pH, although experimental evidence would be required to confirm this.

An alternative reaction mechanism for CPA is known as the nucleophilic pathway, first proposed by Makinen (Makinen *et al.*, 1979), which requires Glu270 to attack the scissile carbonyl of the substrate directly. This nucleophilic attack could not occur in the E204A mutant of LdcB, yet some catalytic turnover is observed (**Figures 4.22 & 4.23**), supporting the idea that LdcB utilises the promoted-water hypothesis instead. The negligible effect of the R120A mutation on catalysis indicates that the bound metal ion is responsible predominantly, if not exclusively, for polarising the carbonyl of the scissile peptide bond. Combining this hypothesis with the structural data described in this thesis supports a reaction mechanism for LdcB that is summarised in **Figure 4.29** (overleaf).

4.6 Summary

This chapter concerns the biochemical characterisation of LdcB and the role of active site residues in ligand binding and catalysis. The identification of key features allowed LdcB to be compared to archetypical carboxypeptidases and a reaction mechanism proposed.



Thin-layer chromatography was used to confirm the designation of both *SpLdcB* and *BsLdcB* as LD-carboxypeptidases, by showing their activity against a synthetic tetrapeptide containing a C-terminal LD amino acid linkage and inactivity against the same tetrapeptide containing a C-terminal LL amino acid linkage. The use of a linked enzyme assay revealed that LdcB enzymes obey Michaelis-Menten kinetics, though the limited availability of the tetrapeptide substrate prevented the determination of accurate kinetic values for *BaLdcB* and *SpLdcB*, and the values obtained for *BsLdcB* are only estimates and likely represent a lower boundary on substrate affinity.

In addition to explaining the observed enzyme kinetics, the structure of the enzyme-product mimic revealed key residues required for the binding of the sugar moiety and first 3 amino acids of the cell wall tetrapeptide, and allowed the prediction of the residues that interact with the 4th amino acid. Furthermore, comparing *SpLdcB* with carboxypeptidase A, VanX and other LAS family enzymes, in addition to data gathered from alanine-substituted mutants, suggests that LdcB uses the promoted-water pathway during catalysis.

Knowledge of the molecular level interactions that govern ligand binding and catalysis provided by this study would permit the design of enzyme inhibitors, which, given the requirement for LdcB for division in *S. pneumoniae* (Barendt *et al.*, 2011), may prove useful as narrow-spectrum antibiotics. Additionally, *SpLdcB* has been shown to be structurally similar to the vancomycin resistance enzymes VanXYc and VanXYg (**Section 3.4.5**), and so interactions that have been identified in *SpLdcB* with M-Tri-Lys(D-Asn) bound likely hold true for VanXYc and VanXYg, opening up the possibility of designing inhibitors against vancomycin resistance.

Chapter 5:

Metal binding in LdcB enzymes

5.1 Introduction

Metal ions are utilised as cofactors in a wide range of enzymes (Waldron and Robinson, 2009). The metal ion can function directly in catalysis, either by forming contacts with the substrate, or by activating a nucleophile for subsequent nucleophilic attack (Christianson and Lipscomb, 1989; Karsisiotis *et al.*, 2014). Alternatively, the large, localised, positive charge of a metal ion can stabilise transition states (Smith *et al.*, 1994; Karsisiotis *et al.*, 2014) or perturb the pK_a of nearby residues (Harris and Turner, 2002), allowing them to participate in catalysis. The displacement of an essential metal ion can therefore reduce or even abolish the activities of metalloproteins (Tamas *et al.*, 2014). Indeed, we have shown in **Chapter 4** that when zinc is removed from *SpLdcB* by incubation with EDTA, catalytic activity is severely abated (**Section 4.3.1**).

In addition to the action of chelators such as EDTA, metal ions can be displaced from enzyme active sites by other metal ions, with the formation of protein:metal complexes with stereo- and/or physico-chemical properties that do not support catalysis (Waldron and Robinson, 2009). For instance, copper has been observed to displace native iron ions from the iron-sulphur cluster enzyme isopropylmalate dehydratase, inactivating the enzyme and inhibiting the biosynthesis of branched-chain amino acids (Macomber and Imlay, 2009). The incorrect metal ion can also often cause conformational changes, contributing to denaturation and aggregation of the protein of interest (Uversky *et al.*, 2001; Lothian *et al.*, 2013).

In this chapter, the interaction of LdcB with zinc, cobalt and copper was assessed (**Sections 5.2 and 5.3**) and the effect of aberrant metal loading on protein structure (**Section 5.4**) and catalysis (**Section 5.5**) was determined.

5.2 Loading non-native metals

ICP-MS has been used already to demonstrate that *SpLdcB* purified with ~50% zinc occupancy, with a small amount (~5%) of nickel (**Section 3.2.2**). In an effort to probe the capacity of *SpLdcB* to bind other *d*-block metals, *SpLdcB* was dialysed overnight against a buffer comprised of 10 mM HEPES pH 8, 100 mM NaCl, 50 mM EDTA to remove all bound metal ions. Following the removal of the EDTA by size exclusion chromatography (HiTrap Desalting, 5 ml [GE Healthcare]), each fraction was diluted

1:10 in Suprapur HNO₃ (Merck) to allow the metal content of the *SpLdcB* fractions to be assessed by ICP-MS (Thermo X-series). The concentrations of metal ions in the fractions were compared to the concentration of *SpLdcB*, as determined by the Bradford method, to assess the metal occupancy of *SpLdcB*. *SpLdcB* was found to be completely free of zinc or nickel (**Figure 5.1**).

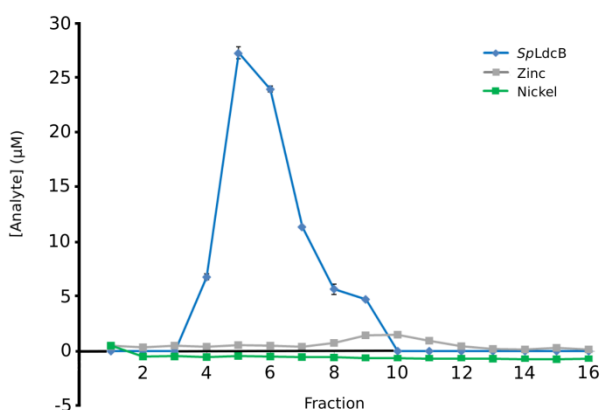


Figure 5.1: Representative metal content analysis of EDTA treated *SpLdcB*. The concentration of zinc and nickel associated with *SpLdcB* after a 16 hour incubation with 50 mM EDTA, as assessed by ICP-MS by measuring [⁶⁶Zn], [⁶⁷Zn] and [⁶⁰Ni]. Protein concentration was determined by the Bradford method. All readings were taken in triplicate and standard deviation shown as error bars.

The metal-free protein was then incubated with a 2-fold molar excess of cobalt chloride or copper chloride to introduce these metals into the active site. Excess metal was removed by size exclusion chromatography (Superdex S75 analytical 10/40 [GE Healthcare]) and each fraction diluted 1:10 in Suprapur HNO₃ (Merck) to allow the metal occupancy of *SpLdcB* to be assessed by ICP-MS (Thermo X-series). Copper was found to bind to *SpLdcB*, though the stoichiometry appears to be greater than 1:1 but less than 2:1 (**Figure 5.2**, overleaf); the peak concentration of copper in the peak fraction of protein is almost twice that of *SpLdcB*, suggesting that a second binding site for copper may be present and partially occupied in solution. Cobalt was determined to bind to *SpLdcB* at a 1:1 stoichiometry, with a large excess of cobalt eluting from the size exclusion column after *SpLdcB* indicating full occupancy of all metal sites (**Figure 5.2**, overleaf).

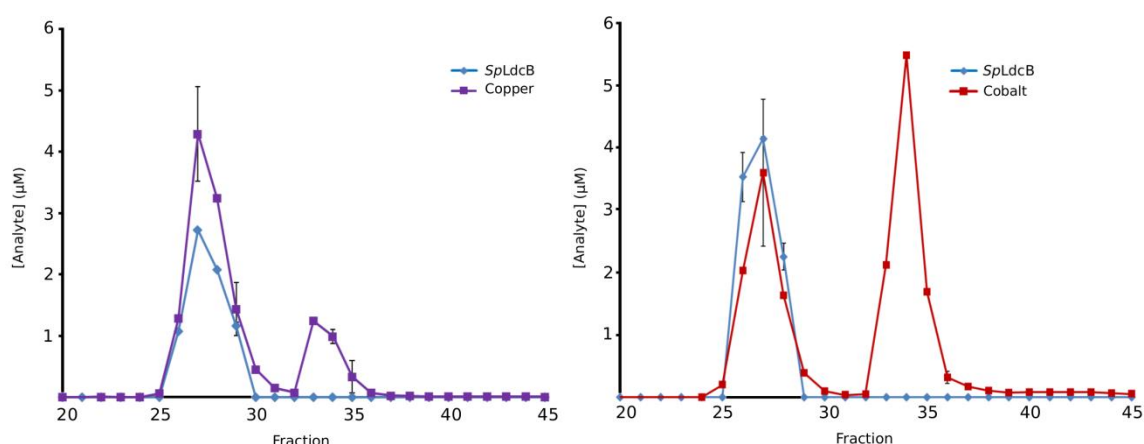


Figure 5.2: Representative metal content analysis of apo-*SpLdcB* following incubation with copper (left) or cobalt (right). The concentration of cobalt and copper associated with *SpLdcB* after a 30 minute incubation with a 2:1 excess of the metal and separation of excess metal by size exclusion chromatography, as assessed by ICP-MS. Protein concentration was determined by the Bradford method. All readings were taken in triplicate and standard deviation shown as error bars.

5.3 The affinity of LdcB for zinc, cobalt and copper

Whilst the Irving-Williams series would suggest LdcB enzymes should purify with copper bound (Irving and Williams, 1953), this was not observed. To investigate why LdcB purified from *E. coli* half occupied with zinc rather than other metals, the affinity of LdcB for zinc, copper and cobalt was assessed. The crystal structure of *SpLdcB* demonstrated that *SpLdcB* possesses multiple metal-binding sites at high metal concentrations, such as those used in crystallisation (**Section 3.4.3**), making *SpLdcB* unsuitable for use in determining the metal affinity of the catalytic metal binding site as any affinity obtained would be an average of all the metal binding sites. Despite the presence of a 4:1 molar excess of zinc in the crystallisation conditions, the crystal structure of *BsLdcB* revealed only a single binding site for metal ions, allowing the affinity of the active site metal binding site to be determined with high accuracy (**Section 3.4.3**).

All native bound metals were removed from *BsLdcB* by dialysis against 10 mM HEPES pH 8, 100 mM NaCl, 50 mM EDTA for 16 hours. EDTA was removed by size exclusion chromatography (Superdex S75 analytical 10/40 [GE Healthcare]) and fractions containing protein were pooled to allow the metal content to be assessed by ICP-MS,

as described previously (**Section 2.4.1**). *BsLdcB* was found to be free of both zinc and nickel.

Increasing concentrations of *BsLdcB* (0 – 120 μM) were added to 4-(2- pyridylazo)resorcinol (PAR) (200 μM) in complex with Zn^{2+} (45 μM), Cu^{2+} (25 μM) or Co^{2+} (10 μM). Uncomplexed PAR absorbs maximally at 415 nm. When PAR forms a complex with a *d*-block metal, the absorption band at 415 nm decreases, and an increase in absorption is seen at higher wavelengths, with the exact wavelength of maximal absorption dependent on the *d*-block metal used (McCall and Fierke, 2000). The absorbance change in PAR upon the addition of increasing *BsLdcB* concentrations was monitored at 500 nm for zinc, or 514 nm for copper and cobalt. The absorbance readings were converted to μM concentrations using previously established extinction coefficients (PAR: Zn^{2+} = 46000 $\text{M}^{-1} \text{cm}^{-1}$ at 500 nm; PAR: Cu^{2+} = 32000 $\text{M}^{-1} \text{cm}^{-1}$ at 514 nm; PAR: Co^{2+} = 50000 $\text{M}^{-1} \text{cm}^{-1}$ at 513 nm (McCall and Fierke, 2000)) and plotted against the concentration of *BsLdcB*. Data were fitted using the following equation, adapted from Badarau and Dennison (2011), using Origin7 (OriginLab):

$$K_d = \frac{[\text{BsLdcB}][\text{PAR}_2\text{M}^{2+}]}{\beta_{\text{PAR}}[\text{BsLdcB} \cdot \text{M}^{2+}][\text{PAR}]^2}$$

in which previously published formation constants (β_{PAR}) of PAR with zinc ($1.26 \times 10^{17} \text{M}^{-2}$ (Zhou *et al.*, 1999)) or copper ($3.2 \times 10^{17} \text{M}^{-2}$ (Turski *et al.*, 2012)) were used. M^{2+} represents the metal ion, either Zn^{2+} , Co^{2+} or Cu^{2+} .

The binding affinities (K_d) of *BsLdcB* for zinc and copper were determined to be $5.61 \times 10^{-17} \text{M}$ (+/- $4.26 \times 10^{-16} \text{M}$) and $2.39 \times 10^{-15} \text{M}$ (+/- $3.31 \times 10^{-14} \text{M}$), respectively (**Figure 5.3**, overleaf), whereas *BsLdcB* was unable to compete with PAR for cobalt (**Figure 5.3**, overleaf).

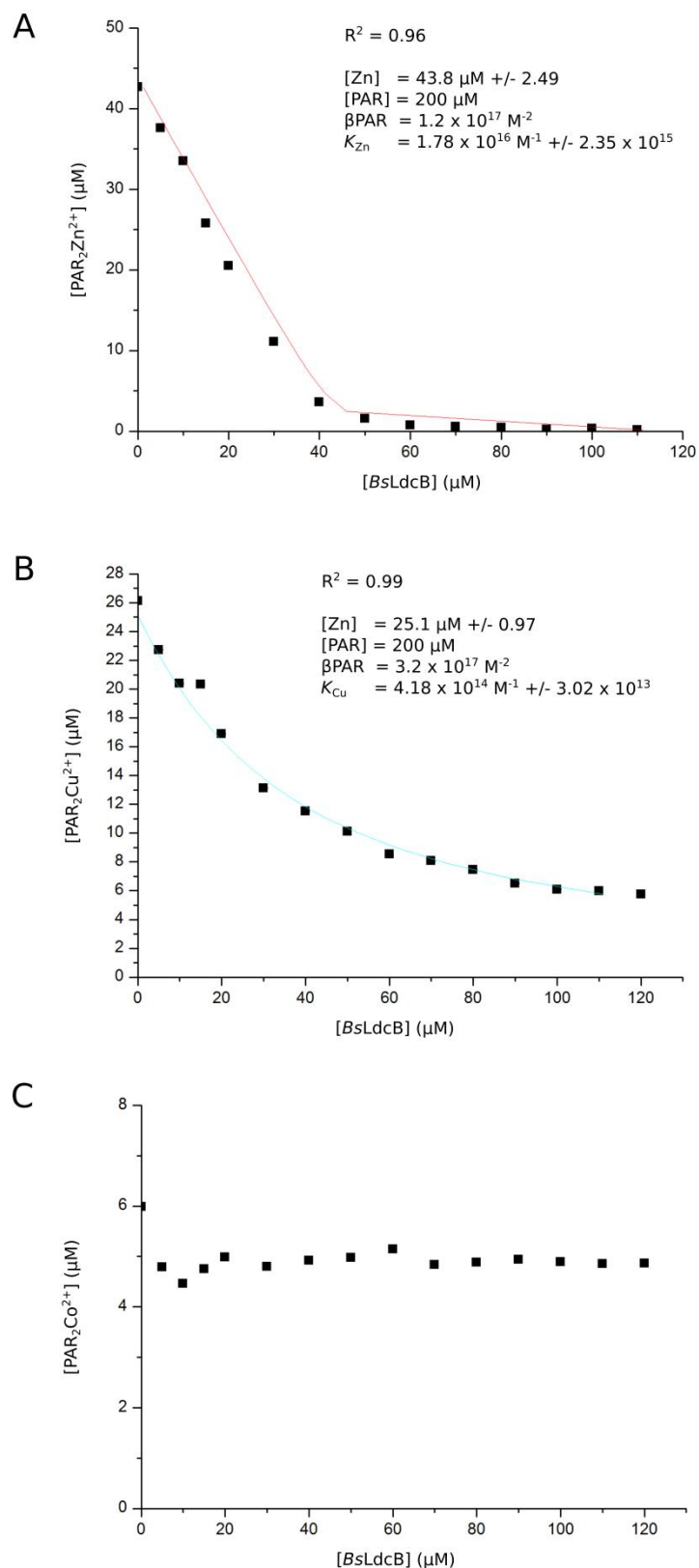


Figure 5.3: Representative data used to determine BsLdcB affinity for (A) Zn^{2+} , (B) Cu^{2+} and (C) Co^{2+} . In all instances increasing concentrations of BsLdcB was added to PAR (200 μM) in complex with Zn^{2+} (45 μM), Cu^{2+} (25 μM) or Co^{2+} (10 μM). Lines show the fit to equation 5.1. K_d values are given in the text.

The toxicity of *d*-block metals, yet their requirement for growth, has resulted in the evolution of complex regulatory systems in all organisms. A simple model would be that a specific metal sensor is triggered when the concentration of 'free' metal ion rises above a key threshold, as determined by the sensor's affinity for that metal (Outten *et al.*, 1999; Munson *et al.*, 2000; Gudipaty and McEvoy, 2014). Should a protein have an affinity for a metal that is below that of the organism's metal sensor, the protein will never encounter the metal ion in a form that it is able to bind, preventing the mis-incorporation of metal ions.

The K_d of the *E. coli* zinc sensor, ZntR, for zinc is 10^{-15} M (Hitomi *et al.*, 2001), 2 orders of magnitude weaker than the affinity of BsLdcB for zinc. BsLdcB is therefore likely able to acquire zinc from the 'free zinc' pool of *E. coli*. With such a great affinity for zinc, it is intriguing that only ~50% of SpLdcB purified with the metal bound (**Section 3.2.2**), possibly suggesting that *E. coli* cannot maintain zinc import at the levels demanded by the overexpression of the heterologous protein. In general, overexpression of any of the LdcB proteins resulted in yields of ~25 mg/l, or $\sim 1 \times 10^{-3}$ mol/l, whilst *E. coli* grown in LB has been shown to have a total zinc content of $\sim 4 \times 10^{-4}$ mol/l (Outten and O'Halloran, 2001). The zinc content of *E. coli* is thus roughly half of that required by the overexpression of LdcB proteins, accounting for only half of purified LdcB containing zinc. One way in which this could be tested would be to add additional zinc salts to the growth media during overexpression and analysing the zinc content of overexpressed LdcB. However, it would be important to limit the amount of zinc added as large excesses are likely to have a toxic effect on the *E. coli*.

The affinity of BsLdcB for copper, 2.39×10^{-15} M, is 2 orders of magnitude weaker than the affinity of BsLdcB for zinc, suggesting that only large excesses of copper would be able to displace zinc from the active site. Within the cell, this would not occur as the *E. coli* copper sensor, CueR, has an affinity for Cu^+ of 10^{-21} M (Changela *et al.*, 2003), resulting in export of copper once the concentration rises above zeptomolar concentrations. It should be noted, however, that the affinity of BsLdcB for copper was determined in aerobic conditions and thus likely corresponds to an affinity for Cu(II), whilst the form of copper encountered in the cytoplasm, a reducing environment (Waldron and Robinson, 2009), would be Cu(I). Nevertheless, unless the affinity of BsLdcB for Cu(I) is 7 or more orders of magnitude stronger than it is for Cu(II), BsLdcB

would never encounter copper in a form that it is able to bind inside the cell. Once on the external leaflet of the cytoplasmic membrane, the physiological location of LdcB enzymes, it is unlikely that copper would be present in sufficient concentrations to effectively displace zinc.

Although the cobalt affinity of *BsLdcB* could not be determined using the competition assay, the absence of cobalt in the purified LdcB protein indicates that the affinity of LdcB for cobalt is weaker than that of the sensor. In *E. coli* the sensor, RcnR, binds cobalt with a nanomolar (10^{-9} M) (Iwig *et al.*, 2008) affinity and thus it can be inferred that *SpLdcB* binds cobalt weaker than 10^{-9} M.

Whilst purified LdcB has been found to purify with small amounts of nickel bound (**Section 3.2.2**), the affinity of *BsLdcB* for nickel was not assessed by PAR, as Ni(PAR)_2 is spectrally silent, making it unsuitable for use in the assay (McCall and Fierke, 2000). It can, however, be inferred that the nickel affinity of *BsLdcB* is less than that of RcnR, which binds nickel with a k_d of at least 2.5×10^{-9} M (Iwig *et al.*, 2008) and acts as a nickel sensor in addition to sensing cobalt, due to nickel only occupying a small fraction of purified LdcB despite ~45% of the protein being metal-free. The small (~5%) nickel content of purified LdcB most likely arises from Ni^+ -NTA affinity purification, with LdcB scavenging nickel from the NTA matrix.

Despite the abundance of data on LAS-family metallopeptidases, there has been relatively few studies to determine the binding affinities of zinc, or other metals, to these proteins. NMB0315, a lysostaphin type peptidase from *Neisseria meningitides*, is the only member of the LAS family to have had its affinity for zinc calculated (Wang *et al.*, 2011). Wang *et al.*, (2011) utilised isothermal titration calorimetry to determine the affinity of NMB0315 for zinc and obtained a K_d value of 1.85×10^{-8} M, much weaker than the femtomolar affinity determined for *BsLdcB*. Inspection of the binding isotherm produced by Wang *et al.*, (2011) provides some explanation for the calculated affinity; the isotherm contains only one data point on the sigmoidal portion of the curve (**Figure 5.4**, overleaf), suggesting that the data may have been overfitted. The sudden increase from 'free' zinc to fully occupied protein observed in the isotherm suggests that the ITC is operating at the limits of its capability, with nanomolar affinities being the tightest binding events measurable by this method, and indicates that the true affinity of NMB0315 for zinc is probably stronger.

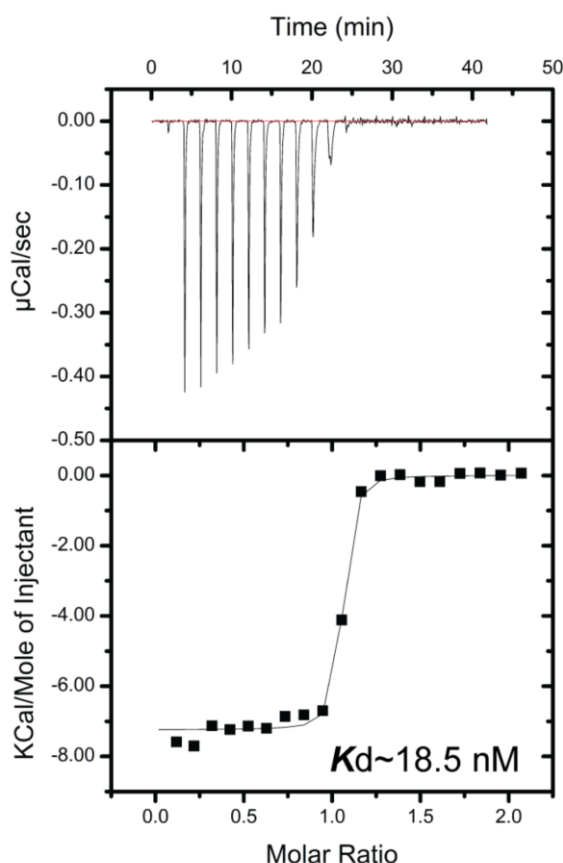


Figure 5.4: ITC isotherm of NMB0315 for zinc taken from Wang *et al.* (2011). The rapid increase in KCal/Mole suggests that the data have been overfitted.

Expanding the search to include non-LAS family enzymes that utilise a His-Asp-His or His-Glu-His motif for zinc binding reveals that the affinities of pancreatic carboxypeptidase A (CPA) for both zinc and cobalt have been calculated (Coleman and Vallee, 1960). Coleman and Vallee (1960) utilised equilibrium dialysis against 100 μM cobalt, with 10 nM radiolabelled cobalt (^{60}Co) added to permit the quantification of bound vs. free cobalt, to determine that CPA binds cobalt with an affinity of 1.5×10^{-6} M. Equilibrium dialysis was then used to determine the affinity of CPA for zinc, by competing both zinc and cobalt for metal-free CPA, with a K_d of 4.7×10^{-9} M obtained for zinc. However, despite the isotope of ^{60}Co allowing for accurate determination of cobalt concentration both within the dialysis buffer and bound to CPA, it is likely that equilibrium was disturbed during measurement, as CPA had to be separated from the dialysis buffer for ^{60}Co quantification, bringing into question the K_d obtained of CPA for cobalt. Indeed, it has been noted that binding affinities of 10^{-6} M are often minimal values for metal binding K_d , rather than true values (Waldron and

Robinson, 2009), suggesting CPA may bind cobalt with a much tighter affinity. Not only would the zinc K_d calculation suffer from similar limitations, the underestimated CPA-cobalt affinity would also affect the calculation, possibly explaining the 6 orders of magnitude difference between CPA zinc affinity and *BsLdcB* zinc affinity, despite both enzymes sharing almost identical zinc binding sites. The only difference is the substitution of Asp142 in *BsLdcB* for Glu72 in CPA.

Searching the literature of all zinc-metalloenzymes reveals a range of binding motifs, though only a select few have had their zinc-affinities calculated. Carbonic anhydrase II coordinates zinc with three histidines (H94, H96 and H119) and a K_d of 4×10^{-12} M (Kiefer *et al.*, 1993). The K_d is increased to 2×10^{-14} M by the addition of an artificial fourth ligand (E199) that occupies the position of the substrate (Ippolito *et al.*, 1995) and which thus greatly reduces activity. The addition of a fourth amino acid ligand is found in nature in the Trx2 subfamily of thioredoxins, in which thioredoxin 2 binds zinc with four cysteine residues and an affinity of 4×10^{-18} M, however, whilst the affinity is comparable to the zinc K_d of *BsLdcB*, the binding motif is not at all similar to LdcB enzymes, which bind with three residues (His, Asp, His).

5.4 Crystal structures of the *SpLdcB*:metal complexes

5.4.1 Crystallisation and structure solution

Although *BsLdcB* proved optimal for determining metal binding affinities, due to its sole metal binding site, the length of time required to grow *BsLdcB* crystals (~6 months, **Section 3.3.3**) makes it an unsuitable protein for structural studies. *SpLdcB* crystals can be obtained in a much shorter amount of time (~1-14 days) and, crucially, the presence of multiple metal binding sites would not adversely affect the experiment.

SpLdcB, in the presence of an excess of cobalt or copper (4:1 molar excess), was subjected to crystallisation trials as described previously (**Sections 3.2 and 3.3**). Crystals of the cobalt containing complexes were observed to grow from the Morpheus (Gorrec, 2009) A4 (0.06 M divalents (MgCl₂ and CaCl₂), 0.1 M Buffer system 1 [imidazole; MES (acid)] pH 6.5, 37.5% MPD_P1K_P3350 [MPD, PEG1000, PEG3350]) and D4 conditions (0.12 M Alcohols [1,6-hexanediol; 1-butanol; 1,2-propandiol (racemic); 2-propanol; 1,4-butanediol; 1,3-propandiol], Buffer system 1 [imidazole;

MES (acid)] pH 6.5, 37.5% MPD_P1K_P3350 [MPD, PEG1000, PEG3350]) over a period of 7 days, while crystals of the copper containing complexes grew from PACT (Newman *et al.*, 2005) C10 (0.2 M MgCl_2 , 0.1 M HEPES pH 7.0, 20% w/v PEG 6000). Crystal optimisation occurred as previously described (**Section 2.6.1**) and was performed under supervision by Ms. Atikah Mohd Sukor. The optimised crystals can be seen in **Figure 5.5**, with the highest quality crystals growing from 0.06 M divalents (MgCl_2 and CaCl_2), 0.1 M Buffer system 1 [imidazole; MES (acid)] pH 6.5, 37.5% MPD_P1K_P3350 [MPD, PEG1000, PEG3350] (*SpLdcB*:cobalt) and from 0.2 M MgCl_2 , 0.1 M HEPES pH 7.0, 20% w/v PEG 6000 (*SpLdcB*:copper).

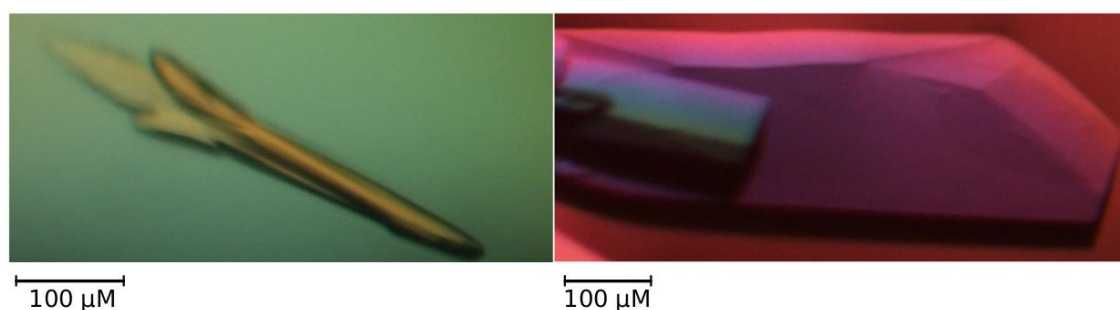


Figure 5.5: Crystal forms of *SpLdcB*:cobalt (left) and *SpLdcB*:copper (right).

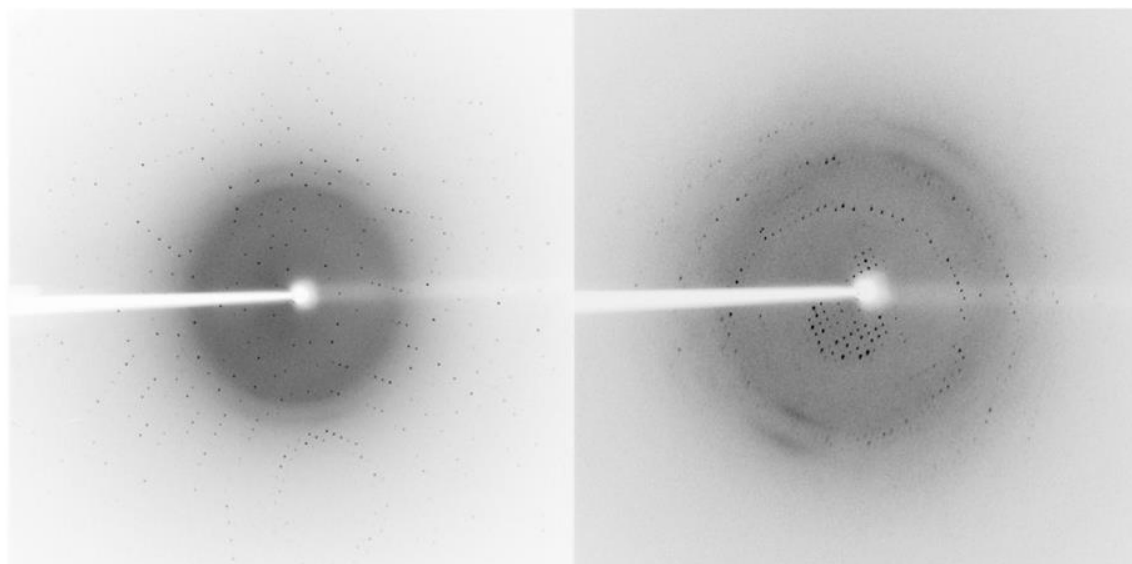


Figure 5.6: An example diffraction pattern of *SpLdcB*:cobalt (left) and *SpLdcB*:copper (right). Cobalt crystals diffracted to $\sim 2 \text{ \AA}$, copper crystals diffracted to $\sim 2.8 \text{ \AA}$.

Crystals were harvested, cryo-protected in paratone-N oil and flash cooled in liquid nitrogen. Diffraction was assessed in-house, with two images being taken at relative phi angles 90° apart. Crystals were ranked based upon spot sharpness, mosaicity and strength of diffraction. Representative diffraction patterns of the best crystals can be found in **Figure 5.6** (previous page). The best crystals were taken to Diamond Light Source for data collection on beamline IO4. Diffraction data sets were collected from the same crystal for both *SpLdcB:cobalt* and *SpLdcB:copper* above and below the theoretical *K* atomic absorption edge (*K*-edge) for the respective metal, with 0.1° oscillations and 0.1 s exposures per image. The total number of images collected was based upon the predicted space group of each crystal to ensure that complete data sets were collected. As with a SAD data collection, the use of incident X-rays at, or just above, the *K*-edge of an element causes an electron transition which results in anomalous scattering and a breakdown in Friedel's law, whereby Friedel pairs no longer scatter with equal amplitude and opposite phases. The breakdown in Friedel's law can be exploited to obtain the phase information required for structure solution. Anomalous difference maps can also be calculated using phases from a solved structure that will generate peaks of difference density that correspond only to the anomalously scattering metal ion.

Diffraction data sets were integrated with XDS (Kabsch, 2010) and reduced with SCALA (Evans, 2005). The *SpLdcB:cobalt* crystal was assigned to space group $P4_3$ with unit cell dimensions of $a = b = 53.6 \text{ \AA}$, $c = 65.0 \text{ \AA}$, based upon the systematic absence of reflections $h+k+l=4n$ and differentiated from the enantiomorphic space group, $P4_1$ based on statistics obtained following molecular replacement and refinement. The *SpLdcB:copper* crystal belonged to spacegroup $P2_1$ with unit cell dimensions of $a = 72.2 \text{ \AA}$, $b = 37.6 \text{ \AA}$, $c = 72.80 \text{ \AA}$, with a β angle of 96.2°. Data collection statistics can be found in **Tables 5.1** and **5.2**.

Table 5.1: Summary of X-ray data collection statistics for *SpLdcB*:cobalt above (High) and below (Low) the K-edge of the ion.

Data collection	<i>SpLdcB</i> :Cobalt High	<i>SpLdcB</i> :Cobalt Low
Space Group	$P4_3$	$P4_3$
Unit cell dimensions		
a, b, c (Å)	53.6, 53.6, 65.0	53.7, 53.7, 65.1
α, β, γ (°)	90, 90, 90	90, 90, 90
Wavelength (Å)	1.605	1.631
Resolution (Å)	41.38 - 1.80 (1.84-1.80)*	41.38 - 1.80 (1.84-1.80)*
Multiplicity	3.2 (3.2)	3.1 (2.5)
Anomalous Multiplicity	1.5 (1.5)	1.5 (1.5)
R_{merge}^a	0.089 (0.212)	0.049 (0.191)
$I/\sigma I$	15.2 (7.8)	15.9 (5.5)
Completeness (%)	99.5 (98.9)	98.8 (94.8)
Anomalous Completeness (%)	86.9 (69.7)	82.0 (21.6)

^a $R_{\text{merge}} = \sum_h \sum_i |I_i(h) - \langle I(h) \rangle| / \sum_h \sum_i I_i(h)$, where $I_i(h)$ and $\langle I(h) \rangle$ are the i th and mean measurement of the intensity of reflection h .

Table 5.2: Summary of X-ray data collection statistics for *SpLdcB*:copper above (High) and below (Low) the K-edge of the ion.

Data collection	<i>SpLdcB</i> :Copper High	<i>SpLdcB</i> :Copper Low
Space Group	$P2_1$	$P2_1$
Unit cell dimensions		
a, b, c (Å)	72.2, 37.6, 72.8	72.4, 37.6, 72.9
α, β, γ (°)	90, 96.2, 90	90, 96.3, 90
Wavelength (Å)	1.494	1.631
Resolution (Å)	48.41 - 2.80 (2.95 - 2.8)	48.45 - 2.80 (2.95 - 2.8)
Multiplicity	3.2 (3.2)	3.1 (3.2)
Anomalous Multiplicity	1.4 (1.8)	1.4 (1.8)
R_{merge}^a	0.039 (0.115)	0.043 (0.078)
$I/\sigma I$	18.3 (8.0)	17.6 (9.9)
Completeness (%)	99.3 (99.6)	96.3 (99.0)
Anomalous Completeness (%)	81.5 (76.3)	73.9 (75.8)

^a $R_{\text{merge}} = \sum_h \sum_i |I_i(h) - \langle I(h) \rangle| / \sum_h \sum_i I_i(h)$, where $I_i(h)$ and $\langle I(h) \rangle$ are the i th and mean measurement of the intensity of reflection h .

Matthews' coefficient calculations (Matthews, 1968) of *SpLdcB*:cobalt predicted that 1 molecule was present per asymmetric unit, with a solvent content of 44.2%, whilst 2 molecules of *SpLdcB*:copper were predicted, with a solvent content of 47.0%. PHASER (McCoy et al., 2007) was used to solve both structures, using the “*SpLdcB*:Cobalt High” and “*SpLdcB*:Copper High” datasets and the co-ordinates of *SpLdcB* as a molecular replacement model. The correctly-positioned protein models were subjected to one round of rigid body refinement in Refmac5 (Winn et al., 2011) before cycles of manual building in COOT and restrained refinement in PHENIX.REFINE (Afonine et al., 2012). Final refinement statistics can be found in **Table 5.3**.

Table 5.3: Summary of X-ray data refinement statistics for *SpLdcB*:cobalt and *SpLdcB*:copper.

	<i>SpLdcB</i> :Cobalt	<i>SpLdcB</i> :Copper
Resolution (Å)	53.6 - 1.8	35.8 – 2.8
No. Reflections	16222 (9700)	14072 (9322)
$R_{\text{work}}/R_{\text{free}}^a$	0.168/0.204	0.252/0.288
No. atoms	-	-
Protein	1434	2733
# proteins/AU	1	2
Ion/Ligand	7/5	11/0
Water	99	12
B factors (Å ²)	-	-
Protein	18.4	36.9/49.9
Ion/Ligand	27.6/15.7	90.3/-
Water	22.4	34.5
R.m.s deviations	-	-
Bond lengths (Å)	0.022	0.009
Bond angles (°)	2.00	1.24
Ramachandran	-	-
favoured (%)	98.9	94.4
allowed (%)	100	99.1

* Highest resolution shell shown in parentheses.

^a $R_{\text{work}}/R_{\text{free}} = \Sigma |F_{\text{p}}^{\text{obs}} - F_{\text{p}}^{\text{calc}}| / \Sigma F_{\text{p}}^{\text{obs}}$, where $F_{\text{p}}^{\text{obs}}$ and $F_{\text{p}}^{\text{calc}}$ are the observed and calculated structure factor amplitudes, respectively.

Following structure solution, anomalous difference maps were calculated. Anomalous difference maps calculated with data collected above the *K*-edge (“High”) revealed difference peaks greater than 10 σ corresponding to the location of the catalytic metal ion (**Figure 5.7**), while difference maps calculated from data collected below the *K*-edge (“Low”) of the respective metal ion were featureless (not shown).

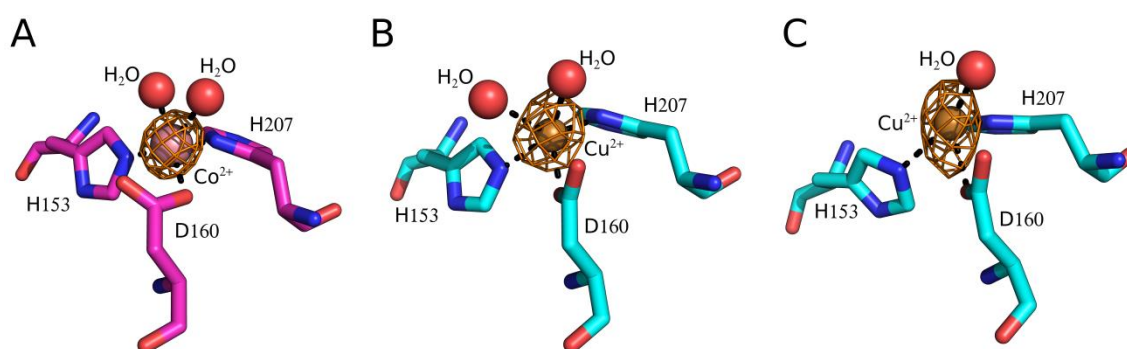


Figure 5.7: Anomalous difference maps. The active sites of cobalt (A) or copper (B = chain A, C = chain B) substituted *SpLdcB* showing the coordination of the metal ion (pink or gold sphere) by water (red sphere) and protein ligands. Anomalous difference maps (orange net) are contoured at at 10 σ . Protein ligands are shown in the stick representation with pink (*SpLdcB*:cobalt) or blue (*SpLdcB*:copper) carbons.

5.4.2 Crystal structures of the *Sp*LdcB:cobalt complex

Overall, the structure of *Sp*LdcB bound to cobalt is indistinguishable from *Sp*LdcB with zinc bound in the closed conformation (core RMSD: 0.57 Å over 181 aligned C α s) (**Figure 5.8**). Closer inspection of the active sites reveals small differences between the structures. As reported previously (**Section 3.4.1**), *Sp*LdcB binds zinc in a tetrahedral coordination sphere; His152, Asp160 and His207 are the liganding amino acids, and an unknown molecule acts as the fourth ligand (**Section 3.4.2**). The *Sp*LdcB:cobalt structure reveals that *Sp*LdcB binds cobalt in a trigonal bipyramidal coordination sphere, with His152, Asp160, His207 and two water molecules acting as ligands (**Figure 5.9**, overleaf).

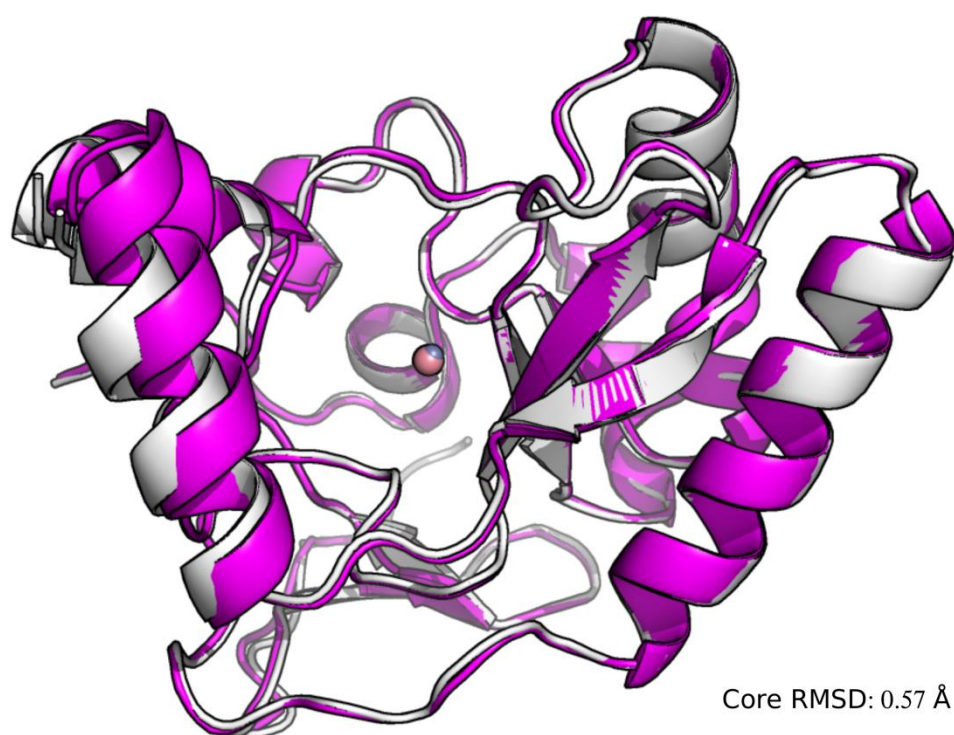


Figure 5.8: *Apo Sp*LdcB with zinc bound superposed onto *Sp*LdcB:cobalt. *Sp*LdcB:zinc is shown as a white cartoon ribbon, *Sp*LdcB:cobalt as a violet cartoon ribbon.

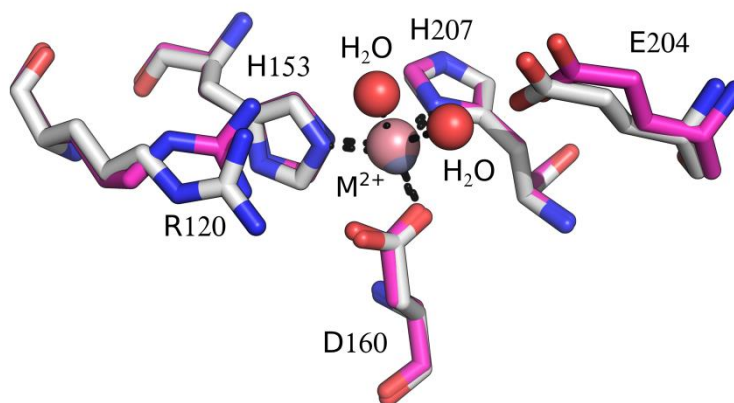


Figure 5.9: Metal binding in *SpLdcB*:cobalt superposed onto the *apo* structure of *SpLdcB* with zinc bound. Cobalt substituted *SpLdcB* coordinates cobalt in a trigonal bipyramidal geometry. His153, Asp160 and His207 contribute contacts from the protein, with solvent molecules completing the 5-coordinate geometry. *SpLdcB*:cobalt carbons are shown in pink, zinc-bound *SpLdcB* carbons are shown in white. The cobalt and zinc ions are indicated by the label M^{2+} , with the cobalt ion shown as a pink sphere, and the zinc ion shown in grey.

The crystal structure of VanX solved previously with bound zinc showed clear tetrahedral coordination of the metal ion (Bussiere *et al.*, 1998). Electronic absorption spectroscopy and extended X-ray absorption fine structure (EXAFS) spectroscopy studies of VanX with cobalt demonstrated that the metal was 5-coordinate (Breece *et al.*, 2005). Breece *et al.*, (2005) argued that cobalt is likely to bind in the same coordination geometry as zinc, owing to the smaller ionic radius, and so postulated that the crystal structure of VanX with zinc bound fails to represent the coordination geometry of VanX:zinc in solution. The ionic radii of *d*-block metals are dependent upon their coordination; whilst cobalt bound in the same geometry as zinc always possesses a smaller ionic radius (tetrahedral Co^{2+} 58 pm vs. Zn^{2+} of 60 pm, trigonal bipyramidal Co^{2+} 67 pm vs. Zn^{2+} of 68 pm (Shannon, 1976)), a 5-coordinate cobalt ion has a much larger ionic radius than 4-coordinate zinc. The expected coordination number can therefore not be determined simply by examining ionic radii as suggested by Breece *et al.*, (2005). Both cobalt and zinc have also been shown to transition between 4- and 5-coordinate without an appreciable energy barrier (Bertini and Claudio, 1994). It is therefore likely that cobalt binds VanX and LdcB in a 5-coordinate manner, while zinc binds 4-coordinate, but that the coordination numbers are fluid. Indeed, a 5-coordinate metal ion is expected for LdcB:zinc during catalysis (Section 4.5).

5.4.3 Crystal structures of the *SpLdcB*:copper complex

Two monomers are present in the asymmetric unit of *SpLdcB*:copper, chains A and B. Both chains are observed to be in the open conformation and are very similar to *SpLdcB* with M-Tri-Lys(D-Asn) bound (**Figure 5.10**) (chain A core RMSD: 1.01 Å over 182 aligned C α s; chain B core RMSD: 0.91 Å over 169 aligned C α s). The open state of the enzyme is perhaps stabilised by a hydrogen bond interaction between Asp116, located on the mobile loop/at the N-terminus of α -helix 4, and Gly67 of a neighbouring chain, with the hydrogen bond only possible in the enzyme's open conformation. However, no other residues could be observed to form crystal contacts, raising the possibility that the enzyme simply happened to be in the open conformation during incorporation into the crystal lattice. Inspection of the active sites reveals that the two chains bind copper in different coordination spheres; copper is found in a trigonal bipyramidal geometry in chain A, whilst it is found tetrahedrally in chain B. The catalytic Glu204 is orientated differently in the two chains; in chain A, Glu204 is positioned to interact with a copper-bound water, mimicking the orientation found in the closed conformation of *SpLdcB* (**Figure 5.11**, overleaf). In chain B, Glu204 adopts the same orientation as when it is interacting with an incoming substrate (**Figure 5.11**, overleaf) (**Section 4.2.3**), despite the lack of electron density indicating a bound ligand.

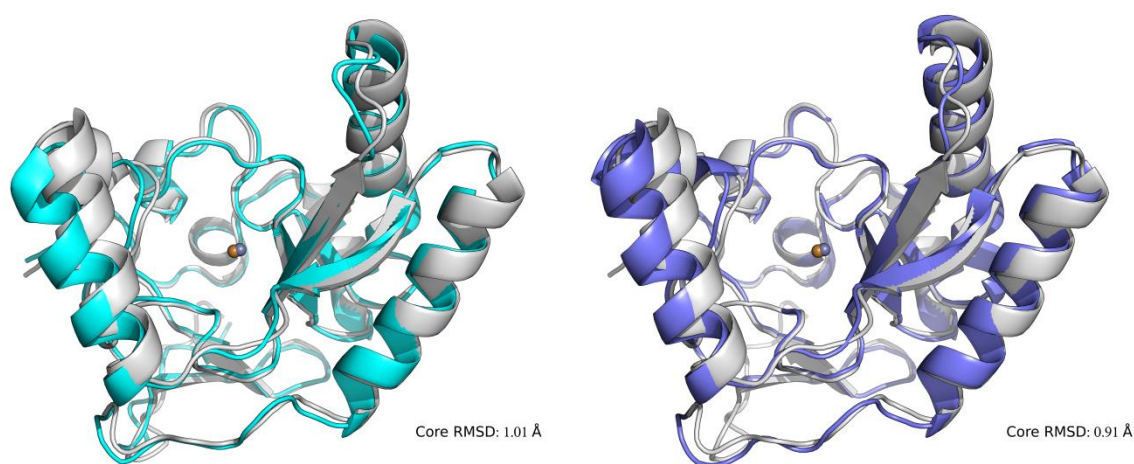


Figure 5.10: Chain A (left) and chain B (right) of *SpLdcB*:copper superposed onto the structure of *SpLdcB* with M-Tri-Lys(D-Asn) bound. *SpLdcB* chain A is shown as light blue cartoon ribbon, chain B is shown as dark blue cartoon ribbon and *SpLdcB*:M-Tri-Lys(D-Asn) is shown as a white cartoon ribbon.

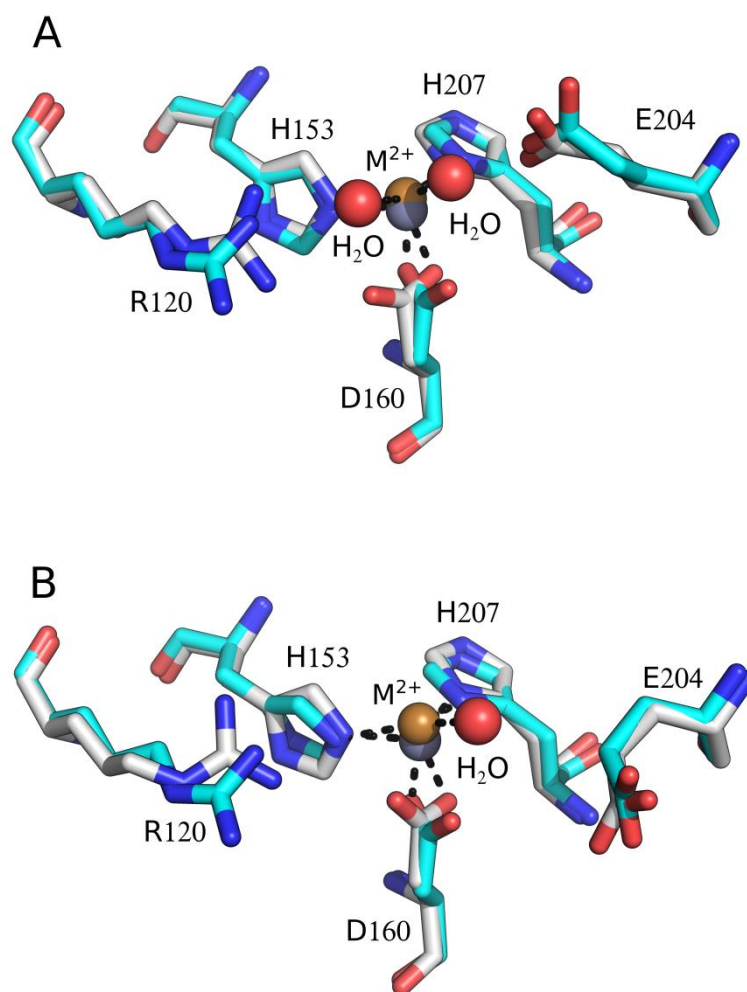


Figure 5.11: The metal binding site of *SpLdcB*:copper, compared to *apo* and ligand-bound *SpLdcB*. Chain A of the *SpLdcB*:copper structure coordinates copper in a trigonal bipyramidal geometry whilst chain B coordinates copper tetrahedrally. Superposing chain A of *SpLdcB*:copper onto *apo SpLdcB* reveals that the catalytic glutamate, E204, is positioned to interact with the nucleophilic water and perform catalysis (A). Superposing chain B of *SpLdcB*:copper onto ligand-bound *SpLdcB* reveals that E204 adopts an orientation required to coordinate the tetrapeptide ligand (B). In both chains A and B, as in *apo* and ligand-bound *SpLdcB*, the metal ion is coordinated by His153, Asp160 and His207. The coordination geometry is completed by two solvent molecules in chain A of *SpLdcB*:copper, one solvent molecule in chain B of *SpLdcB*, an unknown ligand in *apo SpLdcB* (electron density of the unknown ligand not shown) and by a carboxylic oxygen in ligand-bound *SpLdcB* (ligand not shown for clarity). M^{2+} is used to refer to the metal ion, with zinc shown as a grey sphere and copper as a gold sphere. Water molecules are shown as red spheres and amino acids are shown in stick representation with white (*apo SpLdcB*) or blue (*SpLdcB*:copper) carbons.

5.5 The effects of *d*-block metals on catalysis

The substitution of the catalytic ion with other metals has often been used to produce inactive enzymes that remain capable of binding substrates but are unable to process them, trapping enzyme-substrate complexes. However, cobalt-substituted zinc enzymes often retain their catalytic activity and, in some cases, the catalytic rate is increased significantly (Folk and Gladner, 1960; Bertini and Claudio, 1994; Kuzuya and

Inouye, 2001). By contrast, copper has been shown to abolish catalytic activity in a number of zinc enzymes, including thermolysin and β -lactamase II (Hakansson *et al.*, 1994; Hernandez Valladares *et al.*, 2000), although non-negligible activity has been detected in other copper substitutions; Carboxypeptidase A substituted with copper demonstrates k_{cat}/K_m values up to 24% against esters and up to 2.5% towards peptides, compared to the zinc enzyme (Schaffer and Auld, 1986).

The activity of *SpLdcB* substituted with zinc, copper and cobalt was assessed as described previously (**Section 2.5.1.1**), and the reactions were spotted onto a silica TLC plate (Sigma) to allow the products to be resolved (**Figure 5.12**).

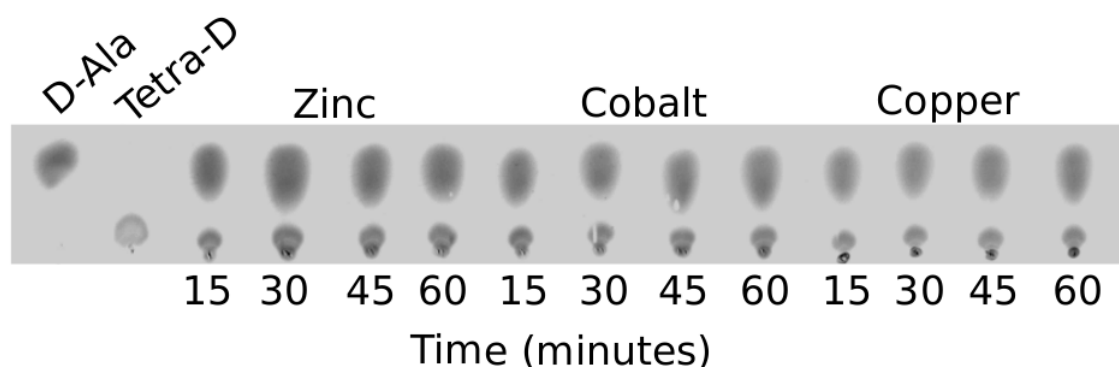


Figure 5.12: Activity of *SpLdcB* loaded with zinc, cobalt or copper against peptidoglycan tetrapeptides. Thin layer chromatography (TLC) of *SpLdcB* loaded with various metals incubated with the tetra-D substrate. D-alanine is less polar than the tetrapeptide and migrates faster in TLC. *SpLdcB* loaded with any of the three *d*-block metals appears to hydrolyse the substrate at similar rates.

No discernible difference in activity was observed between zinc, copper or cobalt-loaded *SpLdcB*, with each reaction running to completion within 15 minutes. These results are compatible with data obtained by Dr. Christine Aldridge, Newcastle University, in which various metal-loaded *SpLdcB* samples were incubated against cell wall material isolated from a $\Delta ldcB$ strain of *S. pneumoniae* (**Figure 5.13**, overleaf). In these experiments *SpLdcB* was incubated against the cell wall substrates for 16 hours prior to HPLC separation of the products, and so this end-point assay cannot reveal differences in rates of activity.

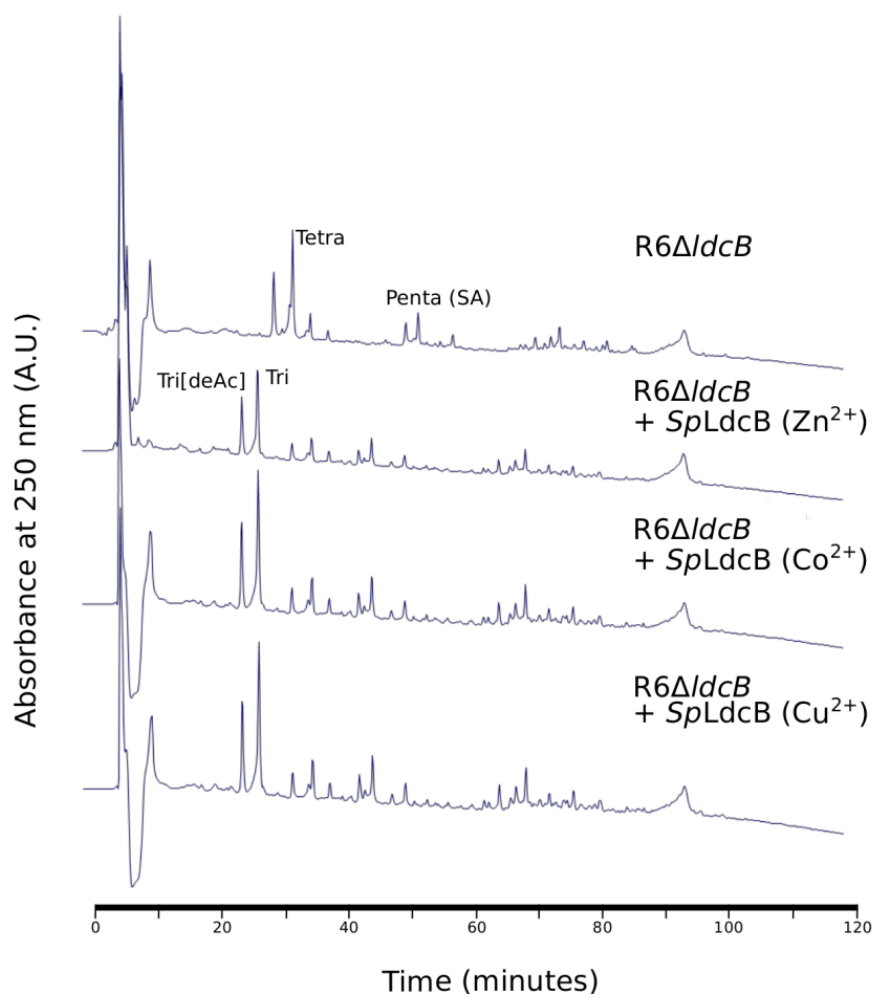


Figure 5.13: HPLC chromatograms of mucopeptides from *S. pneumoniae* strains, obtained after incubating PG with or without *SpLdcB* loaded with zinc, cobalt or copper, followed by digestion with cellosyl and reduction with sodium borohydride. The mucopeptides: Tri, GlcNAc—MurNAc(r)—L-Ala—D-γ-Gln—L-Lys; Tetra, GlcNAc—MurNAc(r)—L-Ala—D-γ-Gln—L-Lys—D-Ala; TetraTri, GlcNAc—MurNAc—L-Ala—D-γ-Gln—L-Lys—D-Ala-L-Lys—D-γ-Gln—L-Ala—MurNAc(r)—GlcNAc; TetraTetra, GlcNAc—MurNAc—L-Ala—D-γ-Gln—L-Lys—(D-Ala)—D-Ala—L-Lys—D-γ-Gln—L-Ala—MurNAc(r)—GlcNAc.

Consistent with the release of D-Ala from Tetra-D (**Figure 5.12**, previous page), cobalt-substituted *SpLdcB* retains full activity against *S. pneumoniae* mucopeptides. As described above, both cobalt and zinc can switch from 4- to 5-coordinate without appreciable energy barriers and their similar ionic radii likely permits their interchangeability (Bertini and Claudio, 1994).

Copper-substituted *SpLdcB* also appears to retain full activity. Although it is possible that the long reaction time in this assay masks differences in activity rates between the various metal-loaded enzymes, the same incubation times were used to assess the E204A and E204H *SpLdcB* mutants (**Section 4.3.2.2**). Clear differences in activity were observed between the wild-type and mutated enzymes within 15 minutes.

Consequently, it would seem that copper does not significantly affect catalysis in comparison to zinc.

Previous comparisons of the active sites of inactive and active copper substituted enzymes have revealed that, in most instances, the copper ion is coordinated in the same geometry as the zinc ion, suggesting that an abnormal coordination sphere is not the cause of inhibition (Bertini and Claudio, 1994). Indeed, the crystal structure of *SpLdcB* with copper bound reveals two coordination geometries; tetrahedral as seen in native *SpLdcB* with zinc bound, and trigonal bipyramidal, which the zinc is predicted to adopt during catalysis (**Section 4.5**). Whilst the catalytic glutamate (E204) is deflected away from the copper ion in chain B of the *SpLdcB*:copper crystal structure, this conformation is also observed in the ligand-bound *SpLdcB* structure, suggesting that the residue is naturally flexible and the presence of copper has not forced the conformational change. Supporting this, chain A of the *SpLdcB*:copper crystal structure reveals that E204 remains located in an ideal position for catalysis, able to interact with one of the copper-coordinating water molecules (**Figure 5.11**).

It is worth noting that previous studies of VanXYc have shown that the addition of either copper or cobalt to purified enzyme causes complete inhibition of dipeptidase activity (Podmore and Reynolds, 2002), which was exploited by Meziane-Cherif *et al.* (2014) to obtain the structure of VanXYc in complex with D-Ala-D-Ala (PDBid: 4OAK). Interestingly, the addition of zinc was found to also inhibit VanXYc (Podmore and Reynolds, 2002), suggesting that excess metal ions may cause adverse conformational changes within VanXYc, rather than directly inhibiting catalysis. The crystal structure of VanXYc with copper bound shows no evidence of adverse conformational changes, however Meziane-Cherif *et al.*, (2014) obtained this structure by soaking a crystal of VanXYc with phenanthroline, a metal chelator, to remove the natively bound zinc, before washing the crystals and soaking in copper, possibly preventing structural changes from occurring. Likewise, Meziane-Cherif *et al.*, (2014) used soaking to introduce D-Ala-D-Ala into the active site of the enzyme, and the kinetics of VanXYc may be radically different within the crystal structure compared to in solution, explaining why the dipeptide was not hydrolysed.

5.6 Summary

This chapter describes metal binding stoichiometries and affinity measurements of *BsLdcB* for zinc, cobalt and copper and presents the crystal structures of *SpLdcB* with copper and cobalt bound, to determine the effect of these metals on protein stability and activity.

Zinc was found to bind to *BsLdcB* with the greatest affinity (5.62×10^{-17} M), whilst copper bound with an affinity 2 orders of magnitude weaker (2.39×10^{-15} M). *BsLdcB* was unable to compete with PAR for cobalt, suggesting a weak affinity but preventing accurate determination of a K_d value. The absence of cobalt from *SpLdcB* and *BsLdcB* purified from *E. coli* suggests that the affinity of LdcB enzymes for this metal is weaker than 10^{-9} M, the affinity of the cobalt sensor RcnR (Iwig *et al.*, 2008). Metal binding affinities of 10^{-6} M or weaker are likely to be significant underestimates (Waldron and Robinson, 2009), suggesting that cobalt may bind to *BsLdcB* with a K_d between 10^{-7} and 10^{-8} M. The PAR assay therefore appears to be unsuitable in determining affinities that are weaker than 10^{-8} M. The affinity of *BsLdcB* for zinc was compared to other zinc-containing proteins and was found to be at the higher-end of zinc affinities, though it was not possible to compare *BsLdcB* with other LAS family enzymes owing to a lack of published literature; the only other LAS family enzyme, NMB0315, with a calculated affinity for zinc (1.85×10^{-8} M) was determined to suffer from experimental errors (**Section 5.3**). Comparing the zinc affinity of *BsLdcB* with carboxypeptidase A (CPA), which share an almost identical, if inverted, active site, suggest that *BsLdcB* binds zinc with 8 orders of magnitude greater affinity than CPA (10^{-17} vs. 10^{-9} M) (Coleman and Vallee, 1960). Given the experiments on CPA were conducted >50 years ago, using radioactive cobalt to determine the amount bound to CPA at equilibrium following dialysis against four different concentrations of cobalt or cobalt-zinc mixtures, it is highly likely that the reported CPA zinc affinity is also underestimated by several orders of magnitude.

The crystal structure of *SpLdcB* with cobalt bound revealed that cobalt binds *SpLdcB* in a manner similar to zinc, although the cobalt ion is coordinated by an extra solvent ligand and so has trigonal bipyramidal geometry. The cobalt structure likely represents the transition state of *SpLdcB*; during catalysis the zinc ion is expected to adopt trigonal bipyramidal geometry to coordinate the reaction intermediate (**Section 4.5**).

Similarly, the crystal structure of *SpLdcB* with copper bound demonstrates that copper is able to bind the enzyme in either a 4- or 5- coordinate geometry, mimicking the coordination spheres of zinc and cobalt, respectively.

It has been demonstrated (**Section 5.5**) that the cobalt and copper substituted forms of *SpLdcB* retain their activity and are not inhibited in any noticeable way. Whilst many cobalt substituted enzymes have been shown to retain, or even show enhanced, activity, it is rare for copper substituted zinc enzymes to continue functioning as efficiently as the enzyme with zinc bound (Folk and Gladner, 1960; Schaffer and Auld, 1986; Bertini and Claudio, 1994; Hakansson *et al.*, 1994; Hernandez Valladares *et al.*, 2000; Kuzuya and Inouye, 2001). In dipeptidyl peptidase III, active site flexibility appears to be a contributory factor in metal-substitution protein activity (Fukasawa *et al.*, 2011); the rigidity of the copper coordination sphere can displace the catalytic residue by a small, but thermodynamically significant, distance to prevent the acid/base/nucleophile from participating in catalysis (Fukasawa *et al.*, 2011). The structure of *SpLdcB* with copper bound clearly demonstrates that the catalytic residues retain similar orientations to those found in the zinc or cobalt bound structures, with the catalytic glutamate, E204, able to be positioned to either perform catalysis or to stabilise incoming ligands or reaction products.

Chapter 6:
Summary and Discussion

6.1 Summary of the major findings of this thesis

The crystal structures of LdcB from *S. pneumoniae* (*SpLdcB*) and *B. subtilis* (*BsLdcB*) are presented in this thesis. Se-Met substituted *SpLdcB* was solved at a resolution of 1.8 Å (PDBid: 4OX5) and *BsLdcB* was solved to 2.0 Å (PDBid: 4OX3). Both structures reveal a single domain, globular protein composed of two sub-domains around a V-shaped active site, at the bottom of which is the catalytic zinc ion (**Section 3.4**). The structures confirmed LdcB as a member of the LAS group of zinc metalloproteases, several members of which are active against peptidoglycan, albeit with different substrate specificities to LdcB. These include the Gly-Gly endopeptidase lysostaphin (Korndörfer *et al.*, 2008) and the DD-/LD- endopeptidase MepA (Marcyjaniak *et al.*, 2004). The structure of *SpLdcB* in complex with a product mimic, M-Tri-Lys(D-Asn), was also solved, which elucidated the molecular basis of peptidoglycan recognition of cell wall modifying members of the LAS family, including the vancomycin resistance enzymes VanY, VanXYg and VanXYc (**Section 4.2**).

This thesis also presents biochemical evidence consistent with LdcB acting as an LD-carboxypeptidase, with the enzyme showing no activity against stem tetrapeptide mimics ending in a canonical LL peptide bond (Ala-D-γ-Gln-Lys-Ala) (**Section 4.3**). Interestingly, LdcB did show minor activity against peptidoglycan pentapeptides ending in a DD-peptide bond (Ala-D-γ-Gln-Lys-D-Ala-D-Ala) (**Section 4.3**), but this activity is unlikely to present itself *in vivo*.

Finally, the requirement for zinc in the active site of the enzyme was assessed, with EDTA-treated LdcB having no activity against substrate (**Section 4.3**), though replacement of the active site zinc with copper or cobalt did not appear to inhibit the enzyme (**Section 5.5**). The affinities of LdcB for zinc, copper and cobalt were determined, which demonstrated that LdcB binds zinc with the highest affinity (**Section 5.3**).

6.2 Other LdcB structures

During the course of this study, another group solved the structure of *SpLdcB*, though they retained the DacB nomenclature for this enzyme (Abdullah *et al.*, 2014). DacB (D-Ala-D-Ala carboxypeptidase B) is the original name designated for *E. coli* PBP4 (Kishida *et al.*, 2006), a DD-endo/DD-carboxypeptidase, and to avoid confusion

Hoyland *et al.*, (2014) suggested that the nomenclature LdcB should be adopted instead when referring to the LD-carboxypeptidase. For the sole purpose of clarity in this section, the structure solved by Abdullah *et al.*, (2014) will be referred to as DacB, whereas the structure reported herein will be referred to as *Sp*LdcB.

Overall, the structure of DacB is almost identical to that of *Sp*LdcB. The two conformations of DacB reported by Abdullah *et al.*, (2014) are consistent with the open and closed conformations observed for *Sp*LdcB (**Section 3.4 and 4.2**). The *apo* form of *Sp*LdcB superposes onto the closed conformation of DacB with a core RMSD of 0.60 Å over 180 aligned C α s, and onto the open conformation with a core RMSD of 0.85 Å over 176 aligned C α s, with the slight increase in RMSD being due to the movement of a loop between Asp167 and Glu171. The ligand-bound form of *Sp*LdcB superposes onto the closed conformation of DacB with a core RMSD of 1.06 Å over 178 aligned C α s, and onto the open conformation with a core RMSD of 1.17 Å over 183 aligned C α s. The major differences between the liganded, open conformation of *Sp*LdcB and the open conformation reported for DacB appear to occur around α -helices 2 and 3 (**Figure 6.1**). These differences are minor – given the small effect on the overall RMSD – and likely result from flexibility in this region of the enzyme. Indeed, the loop connecting α -helix 2 and 3 could not be built in the *apo* structure of *Sp*LdcB (**Section 3.4**) suggesting that local disorder does affect this region of LdcB/DacB.

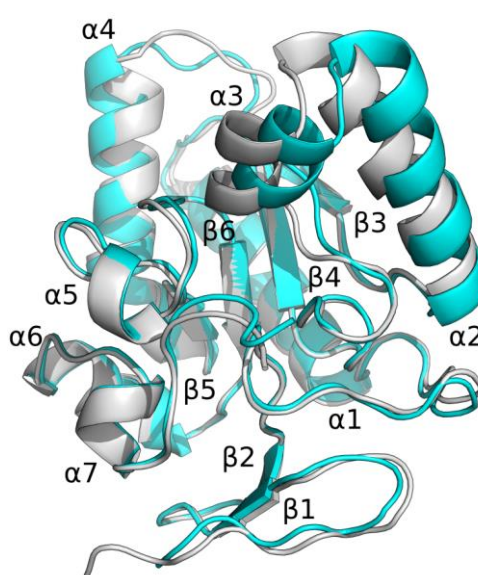


Figure 6.1: Ligand-bound *Sp*LdcB compared to the open conformation of DacB (PDBid: 4D0Y). The major differences between the two structures can be found in α -helices 2 and 3, likely due to disorder of nearby loops.

Neither chain in the DacB structure was found to contain a biologically-relevant ligand, although a phosphate molecule, introduced during purification or crystallisation, coordinates the zinc ion identically to that found in *BsLdcB* (**Section 3.4.2**). Whilst the open conformation of DacB is identical to the ligand-bound form of *SpLdcB*, the catalytic Glu204 remains in the position observed for apo *SpLdcB*. The movement of this residue in the ligand-bound form of *SpLdcB* is therefore not a consequence of the opening of the enzyme, but instead of ligand binding and/or catalysis (**Figure 6.2**).

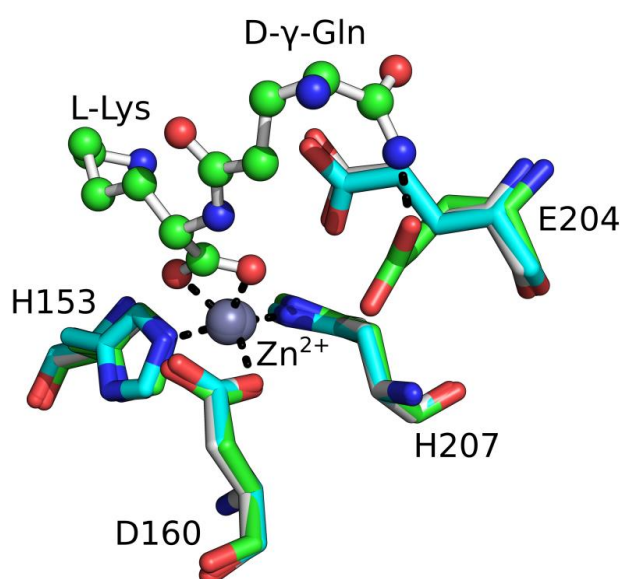


Figure 6.2: The position of Glu204 in the crystal structures of apo *SpLdcB* (PDBid: 40X5), ligand-bound *SpLdcB* (PDBid: 40XD) and the open form of DacB (4D0Y). The catalytic Glu204 in the open conformation of DacB (light blue carbons) is in an identical orientation to Glu204 in the closed form of *SpLdcB* (white carbons). In the open, ligand bound conformation of *SpLdcB* (green carbons) Glu204 moves to interact with D-γ-Gln.

In an attempt to gain insight into substrate recognition by DacB, the structure of VanXYc in complex with D-Ala-D-Ala was superposed onto that of DacB and the D-Ala-D-Ala moiety replaced manually with L-Lys-D-Ala (Abdullah *et al.*, 2014). Abdullah *et al.*, (2014) used the closed conformation of DacB as the model, due to its greater similarity to VanXYc compared to the open conformation. However, the orientation of Trp206 obscures the entrance of the active site and would sterically occlude the full-length tetrapeptide, making the open conformation a better model for this analysis (**Figure 6.3**).

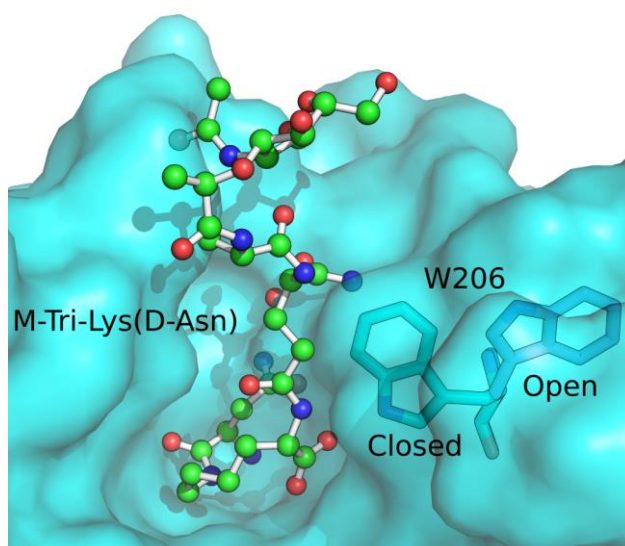


Figure 6.3: The position of Trp206 in open and closed DacB. The closed conformation of Trp206 sterically clashes with the D- γ -Gln of the M-Tri-Lys(D-Asn) ligand (superposed from PDBid: 4OXD), whilst Trp206 in the open conformation allows the ligands to bind to the active site. The M-Tri-Lys(D-Asn) ligand is shown in ball-and-stick representation and the surface of the close conformation of DacB is shown in transparent blue. The two conformations of Trp206 are shown as sticks.

Whilst the contacts between DacB and the D-Ala moiety of the modelled substrate are similar to those between SpLdcB and M-Tri-Lys(D-Asn) (Hoyland *et al.*, 2014), the position of the L-Lys is incompatible with the catalytic properties of the enzyme; the carbonyl oxygen of the L-Lys is not positioned to allow it to complete the tetrahedral coordination geometry of the active site zinc, a requirement of the polarisation of the C=O double bond for nucleophilic attack by an activated water (**Figure 6.4**).

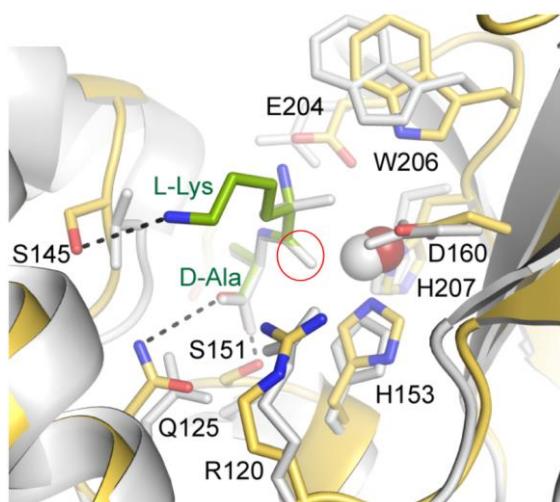


Figure 6.4: The position of the carbonyl oxygen of L-Lys in a model of L-Lys-D-Ala bound to DacB, taken from Abdullah *et al.*, (2014). The carbonyl oxygen (indicated by a red circle) is not coordinating to the metal ion (red or white sphere) when modelled into DacB (gold carbons). The authors superposed DacB onto VanXYg (white carbons) to position the ligand model.

The sidechain of the L-Lys is orientated to form a hydrogen bond with Ser145 in the structure of DacB (**Figure 6.4**, previous page). Whilst this contact is possible in uncrosslinked tetrapeptides, *SpLdcB* is active against both uncrosslinked and crosslinked substrates, suggesting that the sidechain of the L-Lys projects out of the enzyme, along the path proposed by Hoyland *et al.*, (2014). Finally, Abdullah *et al.*, (2014) comment that Trp206 could stabilise the substrate as seen in the structure of VanXYc with D-Ala-D-Ala bound (PDBid: 4MUS) (**Figure 6.5**), in which Trp155 stabilises the amino-terminus of the D-Ala-D-Ala dipeptide. This role is not possible in the conformation adopted by Trp206 in the open, ligand-bound conformation of LdcB (**Section 4.2**).

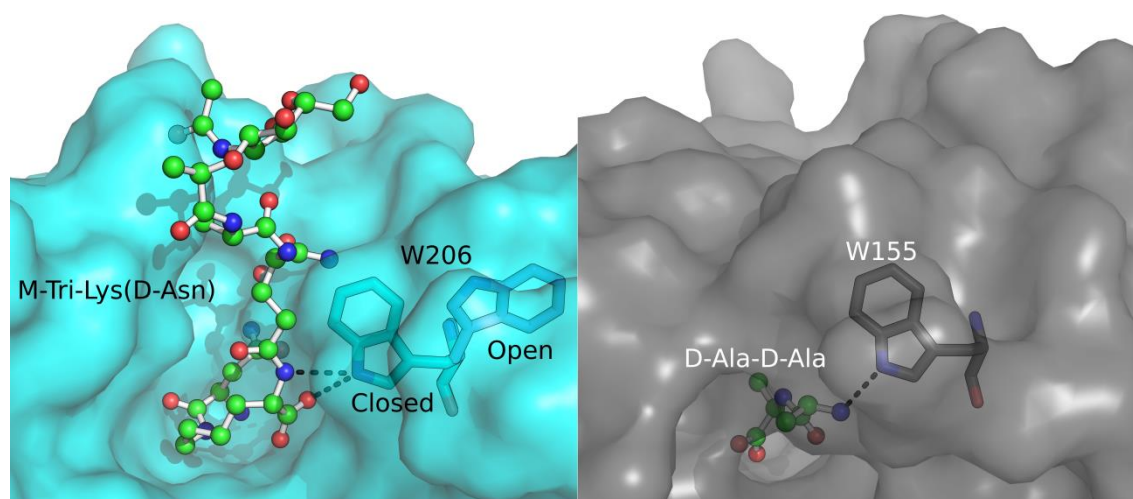


Figure 6.5: The hydrogen bond interaction between Trp206 and the tetrapeptide is not possible in the ligand-bound enzyme conformation. The interactions between Trp206 and the main-chain amino group of the L-Lys in the tetrapeptide ligand in the closed enzyme conformation (left) hypothesised by Abdullah *et al.*, (2014), mimicking the interaction observed between D-Ala-D-Ala and VanXYc by Meziane-Cherif *et al.* (2014) (right), is not possible in the open, ligand bound conformation of LdcB/DacB. The closed conformation of DacB is shown as a blue surface, with the open and closed conformations of Trp206 shown as sticks. The M-Tri-Lys(D-Asn) ligand was obtained by superposing the ligand-bound form of *SpLdcB* (PDBid: 4OXD) onto the closed conformation of DacB, and is shown in ball-and-stick representation. The surface of VanXYc is shown as a transparent grey, with Trp155 shown in stick representation. The D-Ala-D-Ala ligand is shown in ball-and-stick representation.

6.3 The LdcB represents an ancestral form of the vancomycin resistance enzymes

Within the LAS family of which LdcB is a member, LdcB shares most structural and sequence similarity with the vancomycin resistance enzymes VanY, VanXYg and VanXYc (Hoyland *et al.*, 2014; Meziane-Cherif *et al.*, 2014). VanY and VanXYc are active against cell wall pentapeptides, and VanXYg and VanXYc are capable of hydrolysing

D-Ala-D-Ala dipeptides (Meziane-Cherif *et al.*, 2014). Meziane-Cherif *et al.*, (2014) have suggested that LdcB may represent an ancestral form of the vancomycin resistance enzymes, with minor changes to the activity and the specificity of LdcB resulting in the repurposing of the LD-tetracarboxypeptidase into a DD-pentacarboxypeptidase/DD-dipeptidase (Meziane-Cherif *et al.*, 2014). In **Figure 6.6**, a phylogenetic tree compares LdcB to VanY, VanXYg and VanXYc.

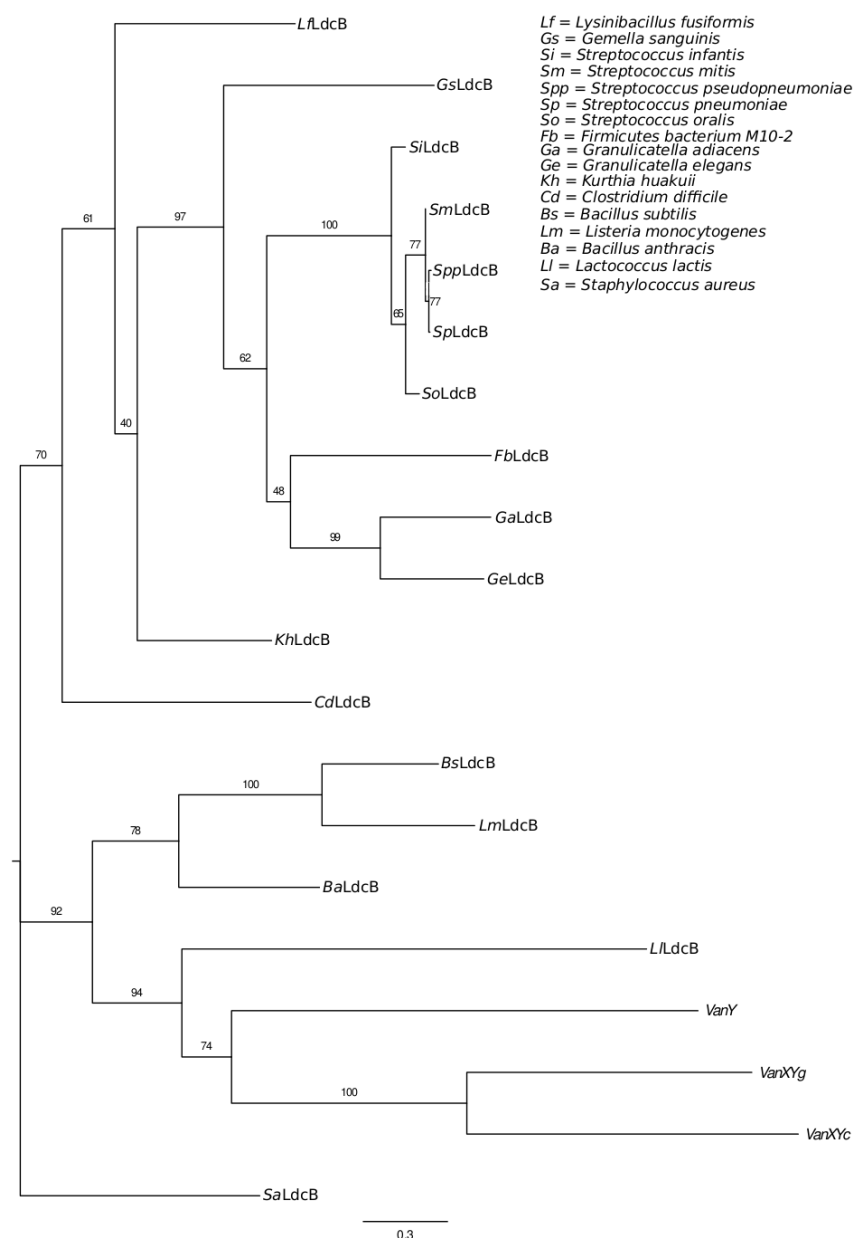


Figure 6.6: Phylogram comparing LdcB sequences with VanY and VanXYg from *Enterococcus faecium*, and VanXYc from *Enterococcus gallinarum*. The VanY and VanXY enzymes occupy a branch of the tree with LdcB from *L. lactis*, with the branch splitting from ancestral LdcB enzymes. Sequence alignments were performed by Muscle (Edgar, 2004), the model was chosen based upon maximum likelihood calculations by the RAXML web-server (Stamatakis, 2006; Stamatakis *et al.*, 2008) and Figtree (<http://tree.bio.ed.ac.uk/software/figtree/>) used to generate the phylogram. Support values for the branches are shown and branch length is related to the number of changes within the protein sequences, weighted by the model chosen by RAXML.

The tree reveals that VanY and the VanXY enzymes appear to be most closely related to LdcB from *Lactococcus lactis* and seem to form a small clade with LdcBs from *B. subtilis*, *B. anthracis* and *L. monocytogenes*, which split from the tree following the divergence of the other LdcB enzymes. It should be noted, however, that the LD-carboxypeptidase activities have only been shown for LdcB from *S. pneumoniae* (Barendt *et al.*, 2011; Abdullah *et al.*, 2014; Hoyland *et al.*, 2014), *L. lactis* (Courtin *et al.*, 2006), *B. subtilis* (Hoyland *et al.*, 2014) and *B. anthracis* (**Section 4.4**), and the other LdcB enzymes have been identified only through sequence similarities. It remains to be seen from whether the other LdcB enzymes possess LD-carboxypeptidase activity.

Comparing the structures of *BsLdcB*, *BaLdcB* and *SpLdcB* might permit the identification of features that confer LD-carboxypeptidase activity, distinct from either DD-carboxypeptidase or D-Ala-D-Ala dipeptidase activity. For example, a loop known as the $\beta 4$ - $\alpha 5$ /bi-substrate selectivity loop determines the substrate selectivity of the VanXY enzymes (Meziane-Cherif *et al.*, 2014). It is believed that in the open conformation of VanXYc the active site is enlarged and the enzyme is capable of binding and hydrolysing the stem pentapeptide, whilst the closed conformation effectively caps the active site and allows for the efficient hydrolysis only of D-Ala-D-Ala dipeptides (Meziane-Cherif *et al.*, 2014). Though LdcB enzymes lack the bi-substrate selectivity loop, they do possess a short, flexible loop ($\beta 4$ - $\alpha 4$) in an equivalent position, the conformation of which changes on tetrapeptide binding (**Section 4.2**) (Hoyland *et al.*, 2014). It is not unreasonable to suggest that a small increase in the length of the $\beta 4$ - $\alpha 4$ loop could result in the formation of the bi-substrate selectivity loop to provide in LdcB enzymes with the ability to hydrolyse D-Ala-D-Ala dipeptides efficiently, an activity that *SpLdcB* possesses only under high substrate concentrations and long incubation times (**Section 4.3.1**). *SpLdcB* also possesses minor activity against pentapeptides (**Section 4.3.1**). A position that could be adopted by the pentapeptide in *SpLdcB* has been modelled (**Section 4.3.1**), and the pentapeptide substrate could make many of the same contacts to the protein as does the tetrapeptide, but with each amino acid of the peptide shifted 1 position upwards. In this position, the MurNAc sugar and the L-alanine project out above the surface of *SpLdcB* and do not make contact with the enzyme (**Figure 6.7**).

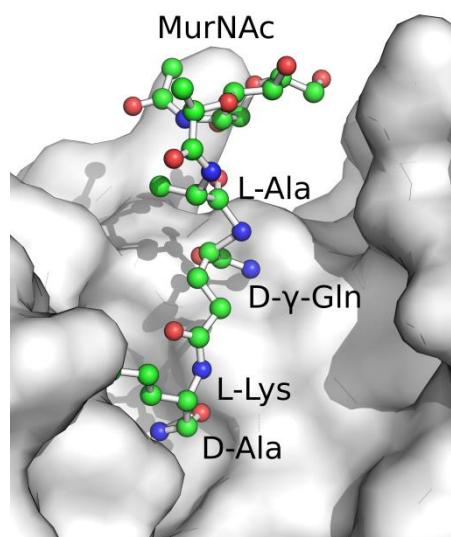


Figure 6.7: The pentapeptide model projecting above the surface of *SpLdcB*. The MurNAc and L-Ala moieties of the pentapeptide model fail to make contact with *SpLdcB*

In the VanXY enzymes, the larger bi-substrate selectivity loop allows Arg119 to contact the O1 of the MurNAc sugar, stabilising the pentapeptide. However in the crystal structures of VanXY published to date, the enzymes are found in their dipeptidase conformation (Meziane-Cherif *et al.*, 2014), and thus any contacts made by the bi-substrate selectivity loop may not be maintained in the presence of a pentapeptide substrate.

Comparing the neighbouring genes of the confirmed LdcB enzymes (*SpLdcB*, *BaLdcB* and *BsLdcB*), using the genome2d webserver (de Jong *et al.*, 2012) to predict operons, suggests that *BsLdcB* and *BaLdcB* do not belong to an operon and are distinct from known peptidoglycan recycling genes. *SpLdcB* is predicted to be part of an operon (operon 0208 in *S. pneumoniae* R6) containing the hypothetical genes spr0552 and spr0553. spr0552 is predicted to code for a hypothetical zinc peptidase of the M42 or M20 family, whilst spr0553 is predicted to code for a hypothetical deaminase enzyme. Without experimental evidence confirming the existence and function of these predicted proteins, hypothesising a function for the operon is difficult, although the presence of the second zinc peptidase may suggest a role in further degrading the stem peptide from a tripeptide to a dipeptide.

6.4 Convergent evolution between peptidases

Although evolutionary distinct, LdcB shares a remarkable similarity with pancreatic carboxypeptidase A (CPA). Whilst the overall structures of the two enzymes are dissimilar, the active sites are essentially mirror-images of one another (**Figure 6.8**), suggesting that similar catalytic mechanisms may be employed. Two catalytic mechanisms have been proposed for CPA, the promoted water pathway and the nucleophilic attack pathway. In the nucleophilic attack pathway, Glu270 of CPA is thought to attack the scissile carbonyl of the substrate to form an acyl-enzyme intermediate, and Arg127 aids to position the substrate. This intermediate then undergoes nucleophilic attack by a zinc-bound water, deacylating the intermediate to result in product formation (Wu *et al.*, 2010). In the promoted water pathway, Glu270 acts as a general base to activate a zinc-bound water for attack on the scissile carbonyl carbon; its polarisation is stimulated by the proximity of the zinc ion and Arg127. This pathway does not result in an acyl-enzyme intermediate, but the substrate forms a tetrahedral intermediate with the catalytic zinc ion instead. Glu270 cleaves the peptide bond of the leaving group by donating a proton (Wu *et al.*, 2010). In SpLdcB, Glu204 and Arg120 are the functional equivalents of CPA Glu270 and Arg127.

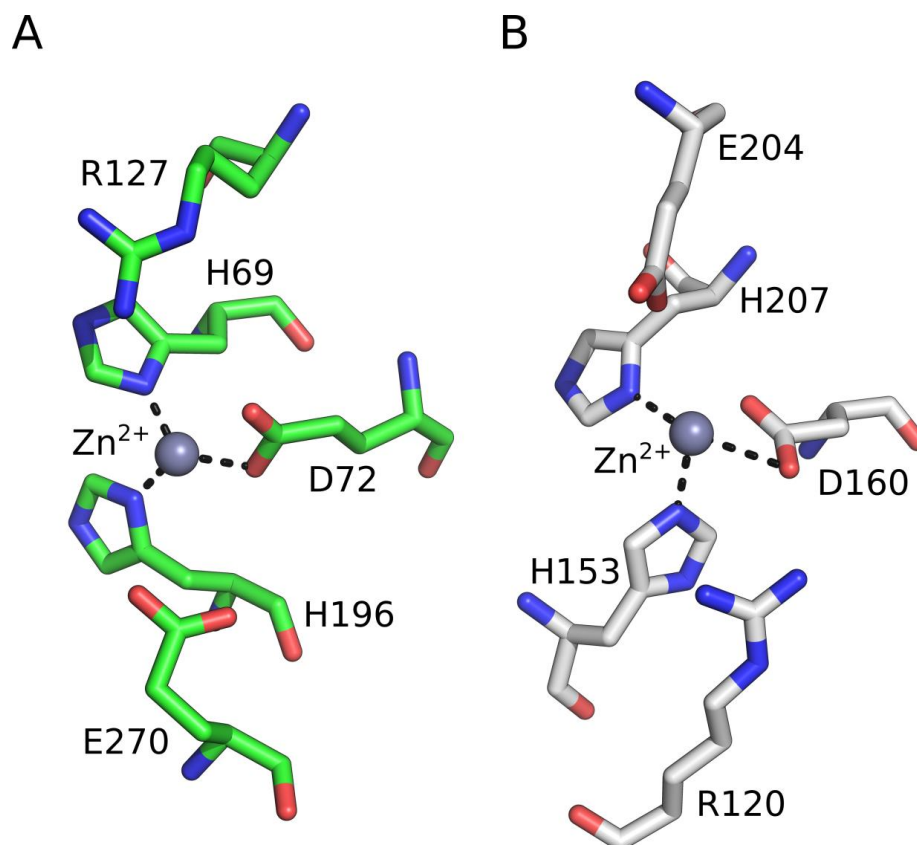


Figure 6.8: The active site of CPA (A) and SpLdcB (B). Despite possessing dissimilar overall structures, the active sites of the two zinc metallopeptidases are remarkably similar. Arg127 and Glu270 of CPA are reversed in SpLdcB (Arg120 and Glu204). CPA carbons are shown in green whilst SpLdcB carbons are shown in white. Both structures are shown in stick representation, with the catalytic zinc ion shown as a grey sphere.

Evidence for the nucleophilic attack mechanism for CPA stems mainly from studying the accumulation of acyl-intermediates at low temperature (Makinen *et al.*, 1979). Ester-containing substrates rather than peptides were used in these studies, and hence the conclusions drawn may not be relevant for proteolysis. However, crystallographic, kinetic and oxygen-exchange studies of CPA support the promoted water pathway (Breslow and Wernick, 1977; Wu *et al.*, 2010). In particular, the oxygen-exchange studies, which monitored the transfer of oxygen-18 from a labelled substrate (i.e. N-benzoylglycine) in the presence of a cocatalyst (i.e. phenylalanine) during the synthesis and subsequent hydrolysis of N-benzoylglycylphenylalanine, noted that the rates of oxygen transfer agreed closely with the rate expected from a direct hydrolysis reaction (e.g. the promoted water pathway), without the formation of any acyl-enzyme intermediates, as would be produced in a nucleophilic attack by a protein amino acid (Breslow and Wernick, 1977). More recently, free-energy simulations of the hydrolysis of esters and peptides have predicted that a lower energy barrier has to be

overcome for the promoted-water pathway than the nucleophilic attack mechanism (Xu *et al.*, 2009; Wu *et al.*, 2010). Some turnover of the tetrapeptide substrate occurs in the Glu204Ala mutant of *SpLdcB* (**Section 4.3.2.2**), which is unlikely to occur should the enzyme utilise the nucleophilic pathway; the nucleophilic pathway requires a carboxylate side-chain, which alanine does not possess but glutamate does, to attack the carbonyl oxygen of the substrate lysine to form an acyl enzyme intermediate. The similarities between the active sites of CPA and LdcB therefore suggest that LdcB is also likely to utilise the promoted water pathway.

In addition to possessing similar active sites, LdcB and CPA are both examples of Koshland's induced fit hypothesis (Koshland, 1958; Christianson and Lipscomb, 1989). Upon binding substrate, CPA undergoes a conformational change in which the backbone of several residues is displaced, and Tyr248 flips from an 'up' position to 'down', allowing it to form hydrogen bonds with the incoming ligand (Steitz *et al.*, 1967). In *SpLdcB*, ligand binding causes a loop between Asp167 and Glu171 to fold into an additional turn in α -helix 4 and the indole ring of Trp206 to move by ~ 9 Å, increasing the size of the active site (**Section 4.2**). The sidechain of Glu204 moves to interact with the D- γ -Gln moiety, though it is unclear if this movement occurs prior to, or post, hydrolysis; the promoted water hypothesis suggests that Glu204 has to re-orientate following catalysis to stabilise the leaving tripeptide, as Glu204 would otherwise be unable of participating as either a general acid or a base. A similar stabilising effect has not been noted for Glu270 of CPA, although the interactions between Tyr248 and the incoming ligand (Steitz *et al.*, 1967) likely act to stabilise the dipeptide in its place. In summary, LdcB and CPA appear to have evolved similar mechanisms for catalysis, despite sharing little sequence identity or overall structural homology, and are therefore excellent examples of convergent evolution.

6.5 The physiological role of LdcB

The physiological effect of *ldcB* deletion appears to differ between bacterial species. For instance, $\Delta ldcB$ *S. pneumoniae* strains develop a compromised cell morphology, with division asymmetry and mislocalisation of division septa (Barendt *et al.*, 2011). The cytoplasm of these cells are less dense than wild-type, suggesting that premature cell lysis occurs (Abdullah *et al.*, 2014). By contrast, deletion of *ldcB* in *B. subtilis* does not result in the appearance of aberrant morphologies during vegetative growth

(Hoyland *et al.*, 2014). LdcB does appear to be necessary for *B. subtilis* growth under D-alanine limiting conditions, however, as $\Delta ldcB$ mutants of a D-alanine auxotroph showed a greater growth-rate dependency on exogenously added D-alanine in comparison to the D-alanine auxotrophic parental strain (Hoyland *et al.*, 2014).

Barendt *et al.*, (2011) have suggested that the compromised cell morphology of $\Delta ldcB$ *S. pneumoniae* is a consequence of the mislocalisation of the penicillin binding proteins (PBPs), the enzymes that facilitate the incorporation of new peptidoglycan into the cell wall during cell elongation or division. This hypothesis is supported by the finding that PBPs localise to their preferred substrate; treating cells with vancomycin (which binds to the terminal D-Ala-D-Ala of the stem pentapeptide to prevent it from participating in crosslinking), mutating the pentapeptide binding domains of PBPs, or preventing nascent peptidoglycan localisation have all been shown to cause delocalisation of PBPs (Morlot *et al.*, 2004; Pinho and Errington, 2005; Potluri *et al.*, 2010; Lages *et al.*, 2013). In addition, this theory implies that stem tetrapeptides, the substrates of LdcB, are a substrate for PBPs; deletion of *ldcB* results in stem tetrapeptides being present over the entire surface of *S. pneumoniae* (Barendt *et al.*, 2011) and coincides with the appearance of aberrant division septa (Barendt *et al.*, 2011). However, the tetrapeptides, which can only function as transpeptidation acceptors, not donors, being a substrate for PBPs is not supported by observations that $\Delta ldcB$ mutants contain fewer crosslinks in their cell walls than wild-type cells (Barendt *et al.*, 2011)

Alternatively, the presence of stem tetrapeptides may localise other cell division proteins. For example StkP, a protein that links cell division and elongation in the Streptococci, has been shown to contain a PASTA domain that is thought to interact with peptidoglycan (Yeats *et al.*, 2002; Beilharz *et al.*, 2012). StkP appears to pause septation (Beilharz *et al.*, 2012), to allow for the peripheral cell wall biosynthesis that is required for cell enlargement prior to division (Higgins and Shockman, 1970; Beilharz *et al.*, 2012). The overexpression of StkP causes the development of small, round cells, and $\Delta stkP$ mutants over-elongate and exhibit multiple division septa (Beilharz *et al.*, 2012). Both StkP phenotypes resemble those of $\Delta ldcB$ *S. pneumoniae*, suggesting a link between StkP and LdcB. By contrast, elongation and division are not explicitly linked in *B. subtilis*; cell elongation occurs prior to and post division rather than during division (Jensen *et al.*, 2005; Schirner and Errington, 2009), thus StkP localisation is not

required to coordinate the two actions, explaining the absence of a morphological phenotype upon *ldcB* deletion.

The growth of *L. lactis*, a close relative of *S. pneumoniae*, has been reported to be unaffected by *ldcB* deletion (Courtin *et al.*, 2006). It is possible that *ldcB* deletion causes a coccus-to-rod transition in *L. lactis*, a naturally-occurring non-detrimental morphological change in this bacterium (Perez-Nunez *et al.*, 2011). The transition occurs when the septation machinery, primarily PBP2X, is halted for a greater amount of time than would occur naturally, whilst peripheral growth continues unabated (Perez-Nunez *et al.*, 2011). As PBP2X co-localises with stem peptide availability (Yeats *et al.*, 2002), it is possible that the *ldcB* deletion causes PBP2X mislocalisation and artificially stalls septation, forcing the cells to become rod-like. Although the hypothesis that a coccus-to-rod transition has occurred could be solved relatively easily by examining $\Delta ldcB$ *L. lactis* under a microscope, linking the transition would require the labelling of PBP2X, and other division proteins, to determine if they localise correctly.

6.6 Concluding remarks and future work

The research undertaken in this thesis has focussed on the LD-carboxypeptidase, LdcB, the biochemistry and 3-D structure of which confirm its designation as an LD-carboxypeptidase. Nevertheless, several avenues of research remain open. First and foremost, the precise mechanism by which LdcB causes morphological changes in *S. pneumoniae*, but not *B. subtilis*, needs to be elucidated. It has been suggested by Barendt *et al.*, (2011) that *ldcB* deletion causes the mislocalisation of PBPs, and $\Delta ldcB$ mutants resemble cells in which StkP, a division coordinator, is either overexpressed or deleted. Clearly, immunofluorescence and/or other methods of determining the localisation patterns cell division proteins (GFP-fusions etc.) need to be employed in conjunction with *ldcB* deletions to help determine the links between LdcB and cell division.

Second, although the liganded SpLdcB structure identified at least some of the amino acids required for peptidoglycan recognition, the path of the L-lysine was not physiologically relevant, nor was the positioning of the D-Asn residue, which occupied the position of the D-Ala leaving group. A path for the lysine to follow when the

enzyme acts on cross-linked substrates has been proposed, but the introduction of bulkier residues (e.g. arginine or tryptophan) and/or the use of longer crosslinks (e.g. a penta-glycine bridge rather than a single D-Asn) is required to validate the path. Additionally, the co-crystallisation of LdcB with the true, tetrapeptide substrate would validate the proposed location of the scissile D-alanine and the interactions that it forms with LdcB.

Third, LdcB might represent a novel, narrow-spectrum antibiotic target. Barendt *et al.*, (2011) demonstrated that LdcB is required for correct septum placement in *S. pneumoniae*, and $\Delta ldcB$ *S. pneumoniae* are much less able than wild-type to colonise and spread in acute pneumonia mouse models, suggesting LdcB may be a viable drug target (Abdullah *et al.*, 2014). The potential design of new antibiotics that target LdcB is aided by the crystal structure of *Sp*LdcB in complex with M-Tri-Lys(D-Asn). In addition, the M-Tri-Lys(D-Asn) ligand bound to *Sp*LdcB may help elucidate the molecular interactions formed between the closely related vancomycin resistance enzymes (VanY, VanXYg and VanXYc) and their stem pentapeptide substrate (**Section 4.2**) and could lead to the development of much needed novel inhibitors to combat the continuing and unchecked march of vancomycin resistance.

Finally, although the work completed herein has been successful in characterising LdcB enzymes, detailed kinetic parameters of the enzymes activity were not obtained. The unavailability of the natural substrate of these enzymes from commercial suppliers is a significant hurdle. The development of an extraction protocol to isolate cell wall tetrapeptides from the peptidoglycan of $\Delta ldcB$ strains would allow the natural substrate of LdcB to be acquired in larger quantities than custom synthesis can provide and would permit a more detailed analysis of the catalytic mechanisms used by LdcB.

References

- Abdullah, M.R., Gutierrez-Fernandez, J., Pribyl, T., Gisch, N., Saleh, M., Rohde, M., Petruschka, L., Burchhardt, G., Schwudke, D., Hermoso, J.A. and Hammerschmidt, S. (2014) 'Structure of the pneumococcal L,d-carboxypeptidase DacB and pathophysiological effects of disabled cell wall hydrolases DacA and DacB', *Molecular Microbiology*, 93(6), pp. 1183-1206.
- Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Echols, N., Headd, J.J., Hung, L.W., Jain, S., Kapral, G.J., Grosse Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R.D., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C. and Zwart, P.H. (2011) 'The Phenix software for automated determination of macromolecular structures', *Methods*, 55(1), pp. 94-106.
- Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H. and Adams, P.D. (2012) 'Towards automated crystallographic structure refinement with phenix.refine', *Acta Crystallographica Section D: Biological Crystallography*, 68(4), pp. 352-367.
- Ainsworth, G.C., Brown, A.M. and Brownlee, G. (1947) 'Aerosporin, an antibiotic produced by *Bacillus aerosporus* Greer', *Nature*, 159(4060), p. 263.
- al-Obeid, S., Billot-Klein, D., van Heijenoort, J., Collatz, E. and Gutmann, L. (1992) 'Replacement of the essential penicillin-binding protein 5 by high-molecular mass PBPs may explain vancomycin-beta-lactam synergy in low-level vancomycin-resistant *Enterococcus faecium* D366', *FEMS Microbiology Letters*, 70(1), pp. 79-84.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990a) 'Basic local alignment search tool', *Journal of Molecular Biology*, 215(3), pp. 403-10.
- Ammann, A.A. (2007) 'Inductively coupled plasma mass spectrometry (ICP MS): a versatile tool', *J Mass Spectrom*, 42(4), pp. 419-27.
- Araoz, R., Anhalt, E., Rene, L., Badet-Denisot, M.A., Courvalin, P. and Badet, B. (2000) 'Mechanism-based inactivation of VanX, a D-alanyl-D-alanine dipeptidase necessary for vancomycin resistance', *Biochemistry*, 39(51), pp. 15971-9.
- Armstrong, J.J., Baddiley, J. and Buchanan, J.G. (1959a) 'Structure of teichoic acid from the walls of *Bacillus subtilis*', *Nature*, 184, pp. 248-9.
- Armstrong, J.J., Baddiley, J., Buchanan, J.G., Davison, A.L., Kelemen, M.V. and Neuhaus, F.C. (1959b) 'Composition of teichoic acids from a number of bacterial walls', *Nature*, 184, pp. 247-8.
- Arthur, M., Reynolds, P. and Courvalin, P. (1996) 'Glycopeptide resistance in enterococci', *Trends in Microbiology*, 4(10), pp. 401-7.
- Atrih, A., Bacher, G., Allmaier, G., Williamson, M.P. and Foster, S.J. (1999) 'Analysis of peptidoglycan structure from vegetative cells of *Bacillus subtilis* 168 and role of PBP 5 in peptidoglycan maturation', *Journal of Bacteriology*, 181(13), pp. 3956-3966.
- B'Hymer, C., Montes-Bayon, M. and Caruso, J.A. (2003) 'Marfey's reagent: Past, present, and future uses of 1-fluoro-2,4-dinitrophenyl-5-L-alanine amide', *Journal of Separation Science*, 26(1-2), pp. 7-19.

- Badarau, A. and Dennison, C. (2011) 'Copper trafficking mechanism of CXXC-containing domains: insight from the pH-dependence of their Cu(I) affinities', *Journal of the American Chemical Society*, 133(9), pp. 2983-8.
- Badet, B., Roise, D. and Walsh, C.T. (1984) 'Inactivation of the *dadB* *Salmonella typhimurium* alanine racemase by D and L isomers of beta-substituted alanines: kinetics, stoichiometry, active site peptide sequencing, and reaction mechanism', *Biochemistry*, 23(22), pp. 5188-94.
- Barendt, S.M., Sham, L.T. and Winkler, M.E. (2011) 'Characterization of mutants deficient in the L,D-carboxypeptidase (DacB) and WalRK (VicRK) Regulon, involved in peptidoglycan maturation of *Streptococcus pneumoniae* serotype 2 strain D39', *Journal of Bacteriology*, 193(9), pp. 2290-2300.
- Battye, T.G., Kontogiannis, L., Johnson, O., Powell, H.R. and Leslie, A.G. (2011) 'iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM', *Acta Crystallogr D Biol Crystallogr*, 67(Pt 4), pp. 271-81.
- Baum, E.Z., Crespo-Carbone, S.M., Foleno, B., Peng, S., Hilliard, J.J., Abbanat, D., Goldschmidt, R. and Bush, K. (2005) 'Identification of a dithiazoline inhibitor of *Escherichia coli* L,D-carboxypeptidase A', *Antimicrobial Agents and Chemotherapy*, 49(11), pp. 4500-7.
- Beilharz, K., Novakova, L., Fadda, D., Branny, P., Massidda, O. and Veening, J.W. (2012) 'Control of cell division in *Streptococcus pneumoniae* by the conserved Ser/Thr protein kinase StkP', *Proceedings of the National Academy of Sciences of the United States of America*, 109(15), pp. E905-13.
- Benson, T.E., Marquardt, J.L., Marquardt, A.C., Etzkorn, F.A. and Walsh, C.T. (1993) 'Overexpression, purification, and mechanistic study of UDP-N-acetylenolpyruvylglucosamine reductase', *Biochemistry*, 32(8), pp. 2024-30.
- Bertini, I. and Claudio, L. (1994) 'The Reaction Pathways of Zinc Enzymes and Related Biological Catalysts', in *Bioinorganic Chemistry*. Mill Valley, CA: University Science Books.
- Bertsche, U., Kast, T., Wolf, B., Fraipont, C., Aarsman, M.E., Kannenberg, K., von Rechenberg, M., Nguyen-Disteche, M., den Blaauwen, T., Holtje, J.V. and Vollmer, W. (2006) 'Interaction between two murein (peptidoglycan) synthases, PBP3 and PBP1B, in *Escherichia coli*', *Molecular Microbiology*, 61(3), pp. 675-90.
- Bhushan, R. and Brückner, H. (2004) 'Marfey's reagent for chiral amino acid analysis: A review', *Amino Acids*, 27(3-4), pp. 231-247.
- Biarrotte-Sorin, S., Hugonnet, J.E., Delfosse, V., Mainardi, J.L., Gutmann, L., Arthur, M. and Mayer, C. (2006) 'Crystal structure of a novel β -lactam-insensitive peptidoglycan transpeptidase', *Journal of Molecular Biology*, 359(3), pp. 533-538.
- Binda, E., Marcone, G.L., Pollegioni, L. and Marinelli, F. (2012) 'Characterization of VanYn, a novel D,D-peptidase/D,D-carboxypeptidase involved in glycopeptide antibiotic resistance in *Nonomuraea* sp. ATCC 39727', *FEBS J*, 279(17), pp. 3203-13.
- Bishop, B., Aricescu, A.R., Harlos, K., O'Callaghan, C.A., Jones, E.Y. and Siebold, C. (2009) 'Structural insights into hedgehog ligand sequestration by the human hedgehog-interacting protein HHIP', *Nature Structural & Molecular Biology*, 16(7), pp. 698-703.

- Bochtler, M., Odintsov, S.G., Marcyjaniak, M. and Sabala, I. (2004) 'Similar active sites in lysostaphins and D-Ala-D-Ala metallopeptidases', *Protein Science*, 13(4), pp. 854-861.
- Bonis, M., Ecobichon, C., Guadagnini, S., Prevost, M.C. and Boneca, I.G. (2010) 'A M23B family metallopeptidase of *Helicobacter pylori* required for cell shape, pole formation and virulence', *Molecular Microbiology*, 78(4), pp. 809-19.
- Boucher, H.W., Talbot, G.H., Bradley, J.S., Edwards, J.E., Gilbert, D., Rice, L.B., Scheld, M., Spellberg, B. and Bartlett, J. (2009) 'Bad Bugs, No Drugs: No ESKAPE! An Update from the Infectious Diseases Society of America', *Clinical Infectious Diseases*, 48(1), pp. 1-12.
- Breece, R.M., Costello, A., Bennett, B., Sigdel, T.K., Matthews, M.L., Tierney, D.L. and Crowder, M.W. (2005) 'A five-coordinate metal center in Co(II)-substituted VanX', *Journal of Biological Chemistry*, 280(12), pp. 11074-81.
- Breslow, R. and Wernick, D.L. (1977) 'Unified picture of mechanisms of catalysis by carboxypeptidase A', *Proceedings of the National Academy of Sciences of the United States of America*, 74(4), pp. 1303-7.
- Bugg, T.D., Wright, G.D., Dutka-Malen, S., Arthur, M., Courvalin, P. and Walsh, C.T. (1991) 'Molecular basis for vancomycin resistance in *Enterococcus faecium* BM4147: biosynthesis of a depsipeptide peptidoglycan precursor by vancomycin resistance proteins VanH and VanA', *Biochemistry*, 30(43), pp. 10408-15.
- Bussiere, D.E., Pratt, S.D., Katz, L., Severin, J.M., Holzman, T. and Park, C.H. (1998) 'The structure of VanX reveals a novel amino-dipeptidase involved in mediating transposon-based vancomycin resistance', *Molecular Cell*, 2(1), pp. 75-84.
- Bzymek, K.P. and Holz, R.C. (2004) 'The catalytic role of glutamate 151 in the leucine aminopeptidase from *Aeromonas proteolytica*', *J Biol Chem*, 279(30), pp. 31018-25.
- Cava, F., Lam, H., de Pedro, M.A. and Waldor, M.K. (2011) 'Emerging knowledge of regulatory roles of D-amino acids in bacteria', *Cell Mol Life Sci*, 68(5), pp. 817-31.
- Changela, A., Chen, K., Xue, Y., Holschen, J., Outten, C.E., O'Halloran, T.V. and Mondragon, A. (2003) 'Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR', *Science*, 301(5638), pp. 1383-7.
- Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S. and Richardson, D.C. (2010) 'MolProbity: all-atom structure validation for macromolecular crystallography', *Acta Crystallogr D Biol Crystallogr*, 66(Pt 1), pp. 12-21.
- Cho, J.H., Kim, D.H., Chung, S.J., Ha, N.C., Oh, B.H. and Yong Choi, K. (2002) 'Insight into the stereochemistry in the inhibition of carboxypeptidase A with N-(hydroxyaminocarbonyl)phenylalanine: binding modes of an enantiomeric pair of the inhibitor to carboxypeptidase A', *Bioorg Med Chem*, 10(6), pp. 2015-22.
- Christianson, D.W. and Lipscomb, W.N. (1989) 'Carboxypeptidase A', *Accounts of Chemical Research*, 22(2), pp. 62-69.
- Claessen, D., Emmins, R., Hamoen, L.W., Daniel, R.A., Errington, J. and Edwards, D.H. (2008) 'Control of the cell elongation-division cycle by shuttling of PBP1 protein in *Bacillus subtilis*', *Molecular Microbiology*, 68(4), pp. 1029-46.

- Cole, R.M. and Hahn, J.J. (1962) 'Cell wall replication in *Streptococcus pyogenes*', *Science*, 135(3505), pp. 722-4.
- Coleman, J.E. and Vallee, B.L. (1960) 'Metallo-carboxypeptidases', *Journal of Biological Chemistry*, 235(2), pp. 390-395.
- Collaborative Computational Project, N. (1994) 'The CCP4 suite: Programs for protein crystallography', *Acta Crystallographica Section D: Biological Crystallography*, 50(5), pp. 760-763.
- Compton, S.J. and Jones, C.G. (1985) 'Mechanism of dye response and interference in the Bradford protein assay', *Anal Biochem*, 151(2), pp. 369-74.
- Costa, K., Bacher, G., Allmaier, G., Dominguez-Bello, M.G., Engstrand, L., Falk, P., de Pedro, M.A. and Garcia-del Portillo, F. (1999) 'The morphological transition of *Helicobacter pylori* cells from spiral to coccoid is preceded by a substantial modification of the cell wall', *Journal of Bacteriology*, 181(12), pp. 3710-5.
- Courtin, P., Miranda, G., Guillot, A., Wessner, F., Mézange, C., Domakova, E., Kulakauskas, S. and Chapot-Chartier, M.P. (2006) 'Peptidoglycan structure analysis of *Lactococcus lactis* reveals the presence of an L,D-carboxypeptidase involved in peptidoglycan maturation', *Journal of Bacteriology*, 188(14), pp. 5293-5298.
- Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004) 'WebLogo: a sequence logo generator', *Genome Research*, 14(6), pp. 1188-90.
- Das, D., Herve, M., Elsliger, M.A., Kadam, R.U., Grant, J.C., Chiu, H.J., Knuth, M.W., Klock, H.E., Miller, M.D., Godzik, A., Lesley, S.A., Deacon, A.M., Mengin-Lecreux, D. and Wilson, I.A. (2013) 'Structure and function of a novel LD-carboxypeptidase involved in peptidoglycan recycling', *Journal of Bacteriology*, 195(24), pp. 5555-66.
- Dauter, Z., Dauter, M. and Dodson, E. (2002) 'Jolly SAD', *Acta Crystallogr D Biol Crystallogr*, 58(Pt 3), pp. 494-506.
- Den Blaauwen, T., Aarsman, M.E., Vischer, N.O. and Nanninga, N. (2003) 'Penicillin-binding protein PBP2 of *Escherichia coli* localizes preferentially in the lateral wall and at mid-cell in comparison with the old cell pole', *Molecular Microbiology*, 47(2), pp. 539-47.
- de Jong, A., Pietersma, H., Cordes, M., Kuipers, O.P. and Kok, J. (2012) 'PePPER: a webserver for prediction of prokaryote promoter elements and regulons', *BMC Genomics*, 13, p. 299.
- Derouaux, A., Wolf, B., Fraipont, C., Breukink, E., Nguyen-Disteche, M. and Terrak, M. (2008) 'The monofunctional glycosyltransferase of *Escherichia coli* localizes to the cell division site and interacts with penicillin-binding protein 3, FtsW, and FtsN', *Journal of Bacteriology*, 190(5), pp. 1831-4.
- Di Berardino, M., Dijkstra, A., Stuber, D., Keck, W. and Gubler, M. (1996) 'The monofunctional glycosyltransferase of *Escherichia coli* is a member of a new class of peptidoglycan-synthesising enzymes', *FEBS Letters*, 392(2), pp. 184-8.
- Di Lallo, G., Fagioli, M., Barionovi, D., Ghelardini, P. and Paolozzi, L. (2003) 'Use of a two-hybrid assay to study the assembly of a complex multicomponent protein machinery: bacterial septosome differentiation', *Microbiology*, 149(Pt 12), pp. 3353-9.

- Dodson, E. (2003) 'Is it jolly SAD?', *Acta Crystallogr D Biol Crystallogr*, 59(Pt 11), pp. 1958-65.
- Double, S. (1997) 'Preparation of selenomethionyl proteins for phase determination', *Methods in Enzymology*, 276, pp. 523-30.
- Doyle, R.J., McDannel, M.L., Streips, U.N., Birdsell, D.C. and Young, F.E. (1974) 'Polyelectrolyte nature of bacterial teichoic acids', *Journal of Bacteriology*, 118(2), pp. 606-15.
- Duncan, K., van Heijenoort, J. and Walsh, C.T. (1990) 'Purification and characterization of the D-alanyl-D-alanine-adding enzyme from *Escherichia coli*', *Biochemistry*, 29(9), pp. 2379-86.
- Edgar, R.C. (2004) 'MUSCLE: multiple sequence alignment with high accuracy and high throughput', *Nucleic Acids Res*, 32(5), pp. 1792-7.
- Emsley, P. and Cowtan, K. (2004) 'Coot: model-building tools for molecular graphics', *Acta Crystallogr D Biol Crystallogr*, 60(Pt 12 Pt 1), pp. 2126-32.
- Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) 'Features and development of Coot', *Acta Crystallographica Section D: Biological Crystallography*, 66(4), pp. 486-501.
- Espaillet, A., Carrasco-Lopez, C., Bernardo-Garcia, N., Pietrosevoli, N., Otero, L.H., Alvarez, L., de Pedro, M.A., Pazos, F., Davis, B.M., Waldor, M.K., Hermoso, J.A. and Cava, F. (2014) 'Structural basis for the broad specificity of a new family of amino-acid racemases', *Acta Crystallogr D Biol Crystallogr*, 70(Pt 1), pp. 79-90.
- Evans, P. (2005) 'Scaling and assessment of data quality', *Acta Crystallographica Section D Biological Crystallography*, 62(1), pp. 72-82.
- Figueiredo, T.A., Sobral, R.G., Ludovice, A.M., de Almeida, J.M.F., Bui, N.K., Vollmer, W., de Lencastre, H. and Tomasz, A. (2012) 'Identification of Genetic Determinants and Enzymes Involved with the Amidation of Glutamic Acid Residues in the Peptidoglycan of *Staphylococcus aureus*', *PLoS Pathog*, 8(1), p. e1002508.
- Fleming, A. (1929) 'On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*', *The British Journal of Experimental Pathology*, 10, pp. 226-236.
- Folk, J.E. and Gladner, J.A. (1960) 'Cobalt activation of carboxypeptidase A', *Journal of Biological Chemistry*, 235, pp. 60-63.
- Fraipont, C., Alexeeva, S., Wolf, B., van der Ploeg, R., Schloesser, M., den Blaauwen, T. and Nguyen-Disteche, M. (2011) 'The integral membrane FtsW protein and peptidoglycan synthase PBP3 form a subcomplex in *Escherichia coli*', *Microbiology*, 157(Pt 1), pp. 251-9.
- Friedrich, E., Vermeulen, J., Biboy, J., Soares, F., Taveirne, M.E., Johnson, J.G., DiRita, V.J., Girardin, S.E., Vollmer, W. and Gaynor, E.C. (2014) 'Peptidoglycan LD-Carboxypeptidase Pgp2 Influences *Campylobacter jejuni* Helical Cell Shape and Pathogenic Properties, and Provides the Substrate for the DL-Carboxypeptidase Pgp1', *Journal of Biological Chemistry*, 289(12), pp. 8007-18.

- Fukasawa, K.M., Hata, T., Ono, Y. and Hirose, J. (2011) 'Metal preferences of zinc-binding motif on metalloproteases', *J Amino Acids*, 2011, p. 574816.
- Garcia-Bustos, J.F., Chait, B.T. and Tomasz, A. (1987) 'Structure of the peptide network of pneumococcal peptidoglycan', *Journal of Biological Chemistry*, 262(32), pp. 15400-5.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D. and Bairoch, A. (2005) 'Protein Identification and Analysis Tools on the ExPASy Server', in Walker, J.M. (ed.) *The Proteomics Protocols Handbook*. Humana Press.
- Georgopapadakou, N., Hammarstrom, S. and Strominger, J.L. (1977) 'Isolation of the penicillin-binding peptide from D-alanine carboxypeptidase of *Bacillus subtilis*', *Proceedings of the National Academy of Sciences of the United States of America*, 74(3), pp. 1009-12.
- Glauner, B., Holtje, J.V. and Schwarz, U. (1988) 'The composition of the murein of *Escherichia coli*', *Journal of Biological Chemistry*, 263(21), pp. 10088-95.
- Goodell, E.W. and Schwarz, U. (1985) 'Release of cell wall peptides into culture medium by exponentially growing *Escherichia coli*', *Journal of Bacteriology*, 162(1), pp. 391-7.
- Gorrec, F. (2009) 'The MORPHEUS protein crystallization screen', *J Appl Crystallogr*, 42(Pt 6), pp. 1035-1042.
- Gross, M., Cramton, S.E., Gotz, F. and Peschel, A. (2001) 'Key role of teichoic acid net charge in *Staphylococcus aureus* colonization of artificial surfaces', *Infection and Immunity*, 69(5), pp. 3423-6.
- Grunbaum, F.A. (1975) 'Remark on the phase problem in crystallography', *Proc Natl Acad Sci USA*, 72(5), pp. 1699-701.
- Gudipaty, S.A. and McEvoy, M.M. (2014) 'The histidine kinase CusS senses silver ions through direct binding by its sensor domain', *Biochimica et Biophysica Acta*, 1844(9), pp. 1656-1661.
- Gunetileke, K.G. and Anwar, R.A. (1968) 'Biosynthesis of uridine diphospho-N-acetylmuramic acid. II. Purification and properties of pyruvate-uridine diphospho-N-acetylglucosamine transferase and characterization of uridine diphospho-N-acetylenopyruvylglucosamine', *Journal of Biological Chemistry*, 243(21), pp. 5770-8.
- Hadley, C (2004) 'Overcoming resistance. Rather than waiting for new drugs, surveillance and education might be more efficient strategies to combat rising antimicrobial resistance', *EMBO reports* 5: 550-552
- Hakansson, K., Wehnert, A. and Liljas, A. (1994) 'X-ray analysis of metal-substituted human carbonic anhydrase II derivatives', *Acta Crystallogr D Biol Crystallogr*, 50(Pt 1), pp. 93-100.
- Hanahan, D. (1983) 'Studies on transformation of *Escherichia coli* with plasmids', *J Mol Biol*, 166(4), pp. 557-80.
- Harris, T.K. and Turner, G.J. (2002) 'Structural basis of perturbed *pKa* values of catalytic groups in enzyme active sites', *IUBMB Life*, 53(2), pp. 85-98.

- Hayashi, R., Bai, Y. and Hata, T. (1975) 'Kinetic studies of carboxypeptidase Y. I. Kinetic parameters for the hydrolysis of synthetic substrates', *J Biochem*, 77, pp. 69-79.
- Hayhurst, E.J., Kailas, L., Hobbs, J.K. and Foster, S.J. (2008) 'Cell wall peptidoglycan architecture in *Bacillus subtilis*', *Proceedings of the National Academy of Sciences*, 105(38), pp. 14603-14608.
- Hernandez Valladares, M., Kiefer, M., Heinz, U., Soto, R.P., Meyer-Klaucke, W., Nolting, H.F., Zeppezauer, M., Galleni, M., Frere, J.M., Rossolini, G.M., Amicosante, G. and Adolph, H.W. (2000) 'Kinetic and spectroscopic characterization of native and metal-substituted beta-lactamase from *Aeromonas hydrophila* AE036', *FEBS Letters*, 467(2-3), pp. 221-5.
- Higgins, M.L. and Shockman, G.D. (1970) 'Model for cell wall growth of *Streptococcus faecalis*', *Journal of Bacteriology*, 101(2), pp. 643-8.
- Hirsch, H.A., Mc, C.C. and Finland, M. (1960) 'Polymyxin B and colistin: activity, resistance and crossresistance *in vitro*', *Proceedings of the Society for Experimental Biology and Medicine*, 103, pp. 338-42.
- Hitomi, Y., Outten, C.E. and O'Halloran, T.V. (2001) 'Extreme zinc-binding thermodynamics of the metal sensor/regulator protein, ZntR', *Journal of the American Chemical Society*, 123(35), pp. 8614-5.
- Holm, L. and Rosenström, P. (2010) 'Dali server: Conservation mapping in 3D', *Nucleic Acids Research*, 38(SUPPL. 2), pp. W545-W549.
- Hoyland, C.N., Aldridge, C., Cleverley, R.M., Duchêne, M., Minasov, G., Onopriyenko, O., Sidiq, K., Stogios, P.J., Anderson, W.F., Daniel, R.A., Savchenko, A., Vollmer, W. and Lewis, R.J. (2014) 'Structure of the LdcB LD-Carboxypeptidase Reveals the Molecular Basis of Peptidoglycan Recognition', *Structure*, 22(7), pp. 949-960.
- Ippolito, J.A., Baird, T.T., Jr., McGee, S.A., Christianson, D.W. and Fierke, C.A. (1995) 'Structure-assisted redesign of a protein-zinc-binding site with femtomolar affinity', *Proceedings of the National Academy of Sciences of the United States of America*, 92(11), pp. 5017-21.
- Irving, H. and Williams, R.J.P. (1953) 'The stability of transition-metal complexes', *Journal of the Chemical Society (Resumed)*, pp. 3192-3210.
- Ito, E. and Strominger, J.L. (1973) 'Enzymatic synthesis of the peptide in bacterial uridine nucleotides. VII. Comparative biochemistry', *Journal of Biological Chemistry*, 248(9), pp. 3131-6.
- Ivey, R.A., Zhang, Y.M., Virga, K.G., Hevener, K., Lee, R.E., Rock, C.O., Jackowski, S. and Park, H.W. (2004) 'The structure of the pantothenate kinase-ADP-pantothenate ternary complex reveals the relationship between the binding sites for substrate, allosteric regulator, and antimetabolites', *Journal of Biological Chemistry*, 279(34), pp. 35622-9.
- Iwig, J.S., Leitch, S., Herbst, R.W., Maroney, M.J. and Chivers, P.T. (2008) 'Ni(II) and Co(II) sensing by *Escherichia coli* RcnR', *Journal of the American Chemical Society*, 130(24), pp. 7592-606.

- Jacobs, C., Joris, B., Jamin, M., Klarsov, K., Van Beeumen, J., Mengin-Lecreulx, D., van Heijenoort, J., Park, J.T., Normark, S. and Frere, J.M. (1995) 'AmpD, essential for both beta-lactamase regulation and cell wall recycling, is a novel cytosolic N-acetylmuramyl-L-alanine amidase', *Molecular Microbiology*, 15(3), pp. 553-9.
- Jefferson, J.R., Hunt, J.B. and Ginsburg, A. (1990) 'Characterization of indo-1 and quin-2 as spectroscopic probes for Zn^{2+} -protein interactions', *Anal Biochem*, 187(2), pp. 328-36.
- Jens, M., Philippe, E., Ima Avalos, V., Philip, K. and Viola, V. (2013) 'Bacterial filamentation accelerates colonization of adhesive spots embedded in biopassive surfaces', *New Journal of Physics*, 15(12), p. 125016.
- Jensen, S.O., Thompson, L.S. and Harry, E.J. (2005) 'Cell division in *Bacillus subtilis*: FtsZ and FtsA association is Z-ring independent, and FtsA is required for efficient midcell Z-Ring assembly', *Journal of Bacteriology*, 187(18), pp. 6536-44.
- Jolly, L., Ferrari, P., Blanot, D., Van Heijenoort, J., Fassy, F. and Mengin-Lecreulx, D. (1999) 'Reaction mechanism of phosphoglucosamine mutase from *Escherichia coli*', *European Journal of Biochemistry*, 262(1), pp. 202-10.
- Kabsch, W. (2010) 'XDS', *Acta Crystallographica Section D: Biological Crystallography*, 66(2), pp. 125-132.
- Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y. and Hattori, M. (2004) 'The KEGG resource for deciphering the genome', *Nucleic Acids Res*, 32(Database issue), pp. D277-80.
- Karimova, G., Pidoux, J., Ullmann, A. and Ladant, D. (1998) 'A bacterial two-hybrid system based on a reconstituted signal transduction pathway', *Proceedings of the National Academy of Sciences of the United States of America*, 95(10), pp. 5752-6.
- Karsisiotis, A.I., Damblon, C.F. and Roberts, G.C. (2014) 'A variety of roles for versatile zinc in metallo-beta-lactamases', *Metallomics*, 6(7), pp. 1181-97.
- Kawai, Y., Marles-Wright, J., Cleverley, R.M., Emmins, R., Ishikawa, S., Kuwano, M., Heinz, N., Bui, N.K., Hoyland, C.N., Ogasawara, N., Lewis, R.J., Vollmer, W., Daniel, R.A. and Errington, J. (2011) 'A widespread family of bacterial cell wall assembly proteins', *EMBO Journal*, 30(24), pp. 4931-41.
- Kelly, S.M., Jess, T.J. and Price, N.C. (2005) 'How to study proteins by circular dichroism', *Biochim Biophys Acta*, 1751(2), pp. 119-39.
- Kelly, S.M. and Price, N.C. (2000) 'The use of circular dichroism in the investigation of protein structure and function', *Curr Protein Pept Sci*, 1(4), pp. 349-84.
- Kiefer, L.L., Krebs, J.F., Paterno, S.A. and Fierke, C.A. (1993) 'Engineering a cysteine ligand into the zinc binding site of human carbonic anhydrase II', *Biochemistry*, 32(38), pp. 9896-900.
- Kilshtain-Vardi, A., Shoham, G. and Goldblum, A. (2002) 'Mechanism of action of zinc proteinases: A MNDO/d/H study of alternative general-acid general-base catalytic pathways for carboxypeptidase-A', *International Journal of Quantum Chemistry*, 88(1), pp. 87-98.

- Kishida, H., Unzai, S., Roper, D.I., Lloyd, A., Park, S.Y. and Tame, J.R. (2006) 'Crystal structure of penicillin binding protein 4 (dacB) from *Escherichia coli*, both in the native form and covalently linked to various antibiotics', *Biochemistry*, 45(3), pp. 783-92.
- Koch, A.L. (2003) 'Were Gram-positive rods the first bacteria?', *Trends in Microbiology*, 11(4), pp. 166-170.
- Kolodkin-Gal, I., Romero, D., Cao, S., Clardy, J., Kolter, R. and Losick, R. (2010) 'D-amino acids trigger biofilm disassembly', *Science*, 328(5978), pp. 627-9.
- Korndörfer, I.P., Kanitz, A., Danzer, J., Zimmer, M., Loessner, M.J. and Skerra, A. (2008) 'Structural analysis of the L-alanoyl-D-glutamate endopeptidase domain of *Listeria bacteriophage* endolysin Ply500 reveals a new member of the LAS peptidase family', *Acta Crystallographica Section D: Biological Crystallography*, 64(6), pp. 644-650.
- Korza, H.J. and Bochtler, M. (2005) '*Pseudomonas aeruginosa* LD-carboxypeptidase, a serine peptidase with a Ser-His-Glu Triad and a nucleophilic elbow', *Journal of Biological Chemistry*, 280(49), pp. 40802-40812.
- Koshland, D.E. (1958) 'Application of a Theory of Enzyme Specificity to Protein Synthesis', *Proc Natl Acad Sci U S A*, 44(2), pp. 98-104.
- Krissinel, E. and Henrick, K. (2007) 'Inference of macromolecular assemblies from crystalline state', *Journal of Molecular Biology*, 372(3), pp. 774-97.
- Kuzuya, K. and Inouye, K. (2001) 'Effects of cobalt-substitution of the active zinc ion in thermolysin on its activity and active-site microenvironment', *J Biochem*, 130(6), pp. 783-8.
- Lages, M.C., Beilharz, K., Morales Angeles, D., Veening, J.W. and Scheffers, D.J. (2013) 'The localization of key *Bacillus subtilis* penicillin binding proteins during cell growth is determined by substrate availability', *Environ Microbiol*, 15(12), pp. 3272-81.
- Lam, H., Oh, D.C., Cava, F., Takacs, C.N., Clardy, J., de Pedro, M.A. and Waldor, M.K. (2009) 'D-amino acids govern stationary phase cell wall remodeling in bacteria', *Science*, 325(5947), pp. 1552-5.
- Lebedev, A.A., Young, P., Isupov, M.N., Moroz, O.V., Vagin, A.A. and Murshudov, G.N. (2012) 'JLigand: a graphical tool for the CCP4 template-restraint library', *Acta Crystallogr D Biol Crystallogr*, 68(Pt 4), pp. 431-40.
- Lee, M., Hesek, D., Llarrull, L.I., Lastochkin, E., Pi, H., Boggess, B. and Mobashery, S. (2013) 'Reactions of all *Escherichia coli* lytic transglycosylases with bacterial cell wall', *Journal of the American Chemical Society*, 135(9), pp. 3311-4.
- Leslie, A.G. (2006) 'The integration of macromolecular diffraction data', *Acta Crystallogr D Biol Crystallogr*, 62(Pt 1), pp. 48-57.
- Liger, D., Masson, A., Blanot, D., van Heijenoort, J. and Parquet, C. (1995) 'Over-production, purification and properties of the uridine-diphosphate-N-acetylmuramate:L-alanine ligase from *Escherichia coli*', *European Journal of Biochemistry*, 230(1), pp. 80-7.

- Lindskog, S. and Coleman, J.E. (1973) 'The catalytic mechanism of carbonic anhydrase', *Proc Natl Acad Sci U S A*, 70(9), pp. 2505-8.
- Lothian, A., Hare, D.J., Grimm, R., Ryan, T.M., Masters, C.L. and Roberts, B.R. (2013) 'Metalloproteomics: principles, challenges and applications to neurodegeneration', *Front Aging Neurosci*, 5, p. 35.
- Lovell, S.C., Davis, I.W., Arendall, W.B., 3rd, de Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S. and Richardson, D.C. (2003) 'Structure validation by C α geometry: phi,psi and C β deviation', *Proteins*, 50(3), pp. 437-50.
- Lovering, A.L., Gretes, M. and Strynadka, N.C. (2008) 'Structural details of the glycosyltransferase step of peptidoglycan assembly', *Current Opinion in Structural Biology*, 18(5), pp. 534-43.
- Lovering, A.L., Safadi, S.S. and Strynadka, N.C.J. (2012) 'Structural perspective of peptidoglycan biosynthesis and assembly', *Annual Review of Biochemistry*, 81, pp. 451-478.
- Lunde, C.S., Rexer, C.H., Hartouni, S.R., Axt, S. and Benton, B.M. (2010) 'Fluorescence microscopy demonstrates enhanced targeting of telavancin to the division septum of *Staphylococcus aureus*', *Antimicrobial Agents and Chemotherapy*, 54(5), pp. 2198-200.
- Macheboeuf, P., Contreras-Martel, C., Job, V., Dideberg, O. and Dessen, A. (2006) 'Penicillin binding proteins: key players in bacterial cell cycle and drug resistance processes', *FEMS Microbiology Reviews*, 30(5), pp. 673-91.
- Macomber, L. and Imlay, J.A. (2009) 'The iron-sulfur clusters of dehydratases are primary intracellular targets of copper toxicity', *Proceedings of the National Academy of Sciences of the United States of America*, 106(20), pp. 8344-9.
- Mäder, U., Schmeisky, A.G., Flórez, L.A. and Stülke, J. (2012) 'SubtiWiki - A comprehensive community resource for the model organism *Bacillus subtilis*', *Nucleic Acids Res*, 40(D1), pp. D1278-D1287.
- Mainardi, J.L., Fourgeaud, M., Hugonnet, J.E., Dubost, L., Brouard, J.P., Ouazzani, J., Rice, L.B., Gutmann, L. and Arthur, M. (2005) 'A novel peptidoglycan cross-linking enzyme for a beta-lactam-resistant transpeptidation pathway', *Journal of Biological Chemistry*, 280(46), pp. 38146-52.
- Makinen, M.W., Kuo, L.C., Dymowski, J.J. and Jaffer, S. (1979) 'Catalytic role of the metal ion of carboxypeptidase A in ester hydrolysis', *Journal of Biological Chemistry*, 254(2), pp. 356-66.
- Marcyjaniak, M., Odintsov, S.G., Sabala, I. and Bochtler, M. (2004) 'Peptidoglycan amidase MepA is a LAS metallopeptidase', *Journal of Biological Chemistry*, 279(42), pp. 43982-43989.
- Matthews, B.W. (1968) 'Solvent content of protein crystals', *Journal of Molecular Biology*, 33(2), pp. 491-7.
- McCall, K.A. and Fierke, C.A. (2000) 'Colorimetric and fluorimetric assays to quantitate micromolar concentrations of transition metals', *Analytical Biochemistry*, 284(2), pp. 307-15.

- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. (2007) 'Phaser crystallographic software', *Journal of Applied Crystallography*, 40(4), pp. 658-674.
- McKessar, S.J., Berry, A.M., Bell, J.M., Turnidge, J.D. and Paton, J.C. (2000) 'Genetic characterization of vanG, a novel vancomycin resistance locus of *Enterococcus faecalis*', *Antimicrobial Agents and Chemotherapy*, 44(11), pp. 3224-8.
- McKevitt, M.T., Bryant, K.M., Shakir, S.M., Larabee, J.L., Blanke, S.R., Lovchik, J., Lyons, C.R. and Ballard, J.D. (2007) 'Effects of endogenous D-alanine synthesis and autoinhibition of *Bacillus anthracis* germination on in vitro and in vivo infections', *Infection and Immunity*, 75(12), pp. 5726-34.
- Mengin-Lecreulx, D., Texier, L., Rousseau, M. and van Heijenoort, J. (1991) 'The murG gene of *Escherichia coli* codes for the UDP-N-acetylglucosamine: N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase involved in the membrane steps of peptidoglycan synthesis', *Journal of Bacteriology*, 173(15), pp. 4625-36.
- Mengin-Lecreulx, D. and van Heijenoort, J. (1993) 'Identification of the glmU gene encoding N-acetylglucosamine-1-phosphate uridylyltransferase in *Escherichia coli*', *Journal of Bacteriology*, 175(19), pp. 6150-7.
- Mercer, K.L. and Weiss, D.S. (2002) 'The *Escherichia coli* cell division protein FtsW is required to recruit its cognate transpeptidase, FtsI (PBP3), to the division site', *Journal of Bacteriology*, 184(4), pp. 904-12.
- Meziane-Cherif, D., Stogios, P.J., Evdokimova, E., Savchenko, A. and Courvalin, P. (2014) 'Structural basis for the evolution of vancomycin resistance D,D-peptidases', *Proceedings of the National Academy of Sciences of the United States of America*, 111(16), pp. 5872-7.
- Mitchell, P. and Moyle, J. (1957) 'Autolytic release and osmotic properties of protoplasts from *Staphylococcus aureus*', *Journal of General Microbiology*, 16(1), pp. 184-94.
- Mock, W.L. and Tsay, J.T. (1988) 'pK values for active site residues of carboxypeptidase A', *J Biol Chem*, 263(18), pp. 8635-41.
- Mohammadi, T., van Dam, V., Sijbrandi, R., Vernet, T., Zapun, A., Bouhss, A., Diepeveen-de Bruin, M., Nguyen-Disteche, M., de Kruijff, B. and Breukink, E. (2011) 'Identification of FtsW as a transporter of lipid-linked cell wall precursors across the membrane', *EMBO Journal*, 30(8), pp. 1425-32.
- Morlot, C., Noirclerc-Savoye, M., Zapun, A., Dideberg, O. and Vernet, T. (2004) 'The D,D-carboxypeptidase PBP3 organizes the division process of *Streptococcus pneumoniae*', *Molecular Microbiology*, 51(6), pp. 1641-8.
- Morlot, C., Pernot, L., Le Gouellec, A., Di Guilmi, A.M., Vernet, T., Dideberg, O. and Dessen, A. (2005) 'Crystal structure of a peptidoglycan synthesis regulatory factor (PBP3) from *Streptococcus pneumoniae*', *Journal of Biological Chemistry*, 280(16), pp. 15984-15991.

- Morlot, C., Zapun, A., Dideberg, O. and Vernet, T. (2003) 'Growth and division of *Streptococcus pneumoniae*: localization of the high molecular weight penicillin-binding proteins during the cell cycle', *Molecular Microbiology*, 50(3), pp. 845-55.
- Munson, G.P., Lam, D.L., Outten, F.W. and O'Halloran, T.V. (2000) 'Identification of a copper-responsive two-component system on the chromosome of *Escherichia coli* K-12', *Journal of Bacteriology*, 182(20), pp. 5864-71.
- Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) 'Refinement of macromolecular structures by the maximum-likelihood method', *Acta Crystallogr D Biol Crystallogr*, 53(Pt 3), pp. 240-55.
- Nagarajan, R. (1991) 'Antibacterial activities and modes of action of vancomycin and related glycopeptides', *Antimicrobial Agents and Chemotherapy*, 35(4), pp. 605-9.
- Neuhaus, F.C. and Baddiley, J. (2003) 'A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in Gram-positive bacteria', *Microbiology and Molecular Biology Reviews*, 67(4), pp. 686-723.
- Newman, J., Egan, D., Walter, T.S., Meged, R., Berry, I., Ben Jelloul, M., Sussman, J.L., Stuart, D.I. and Perrakis, A. (2005) 'Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: the PACT/JCSG+ strategy', *Acta Crystallogr D Biol Crystallogr*, 61(Pt 10), pp. 1426-31.
- Nystrom, T. (2004) 'Stationary-phase physiology', *Annu Rev Microbiol*, 58, pp. 161-81.
- Odintsov, S.G., Sabala, I., Marcyjaniak, M. and Bochtler, M. (2004) 'Latent LytM at 1.3Å resolution', *Journal of Molecular Biology*, 335(3), pp. 775-85.
- Oswald, C., Smits, S.H., Bremer, E. and Schmitt, L. (2008) 'Microseeding - a powerful tool for crystallizing proteins complexed with hydrolyzable substrates', *Int J Mol Sci*, 9(7), pp. 1131-41.
- Otero, L.H., Rojas-Altuve, A., Llarrull, L.I., Carrasco-Lopez, C., Kumarasiri, M., Lastochkin, E., Fishovitz, J., Dawley, M., Heseck, D., Lee, M., Johnson, J.W., Fisher, J.F., Chang, M., Mobashery, S. and Hermoso, J.A. (2013) 'How allosteric control of *Staphylococcus aureus* penicillin binding protein 2a enables methicillin resistance and physiological function', *Proc Natl Acad Sci U S A*, 110(42), pp. 16808-13.
- Outten, C.E. and O'Halloran, T.V. (2001) 'Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis', *Science*, 292(5526), pp. 2488-92.
- Outten, C.E., Outten, F.W. and O'Halloran, T.V. (1999) 'DNA distortion mechanism for transcriptional activation by ZntR, a Zn(II)-responsive MerR homologue in *Escherichia coli*', *Journal of Biological Chemistry*, 274(53), pp. 37517-24.
- Painter, J. and Merritt, E.A. (2006) 'TLSMD web server for the generation of multi-group TLS models', *Journal of Applied Crystallography*, 39(1), pp. 109-111.
- Pape, T. and Schneider, T.R. (2004) 'HKL2MAP: a graphical user interface for phasing with SHELX programs', *J. Appl. Cryst.*, 37, pp. 843-844.

- Pastoret, S., Fraipont, C., den Blaauwen, T., Wolf, B., Aarsman, M.E., Piette, A., Thomas, A., Brasseur, R. and Nguyen-Disteche, M. (2004) 'Functional analysis of the cell division protein FtsW of *Escherichia coli*', *Journal of Bacteriology*, 186(24), pp. 8370-9.
- Paterson, D.L. and Bonomo, R.A. (2005) 'Extended-spectrum beta-lactamases: a clinical update', *Clinical Microbiology Reviews*, 18(4), pp. 657-86.
- Perez-Nunez, D., Briandet, R., David, B., Gautier, C., Renault, P., Hallet, B., Hols, P., Carballido-Lopez, R. and Guedon, E. (2011) 'A new morphogenesis pathway in bacteria: unbalanced activity of cell wall synthesis machineries leads to coccus-to-rod transition and filamentation in ovococci', *Molecular Microbiology*, 79(3), pp. 759-71.
- Pinho, M.G. and Errington, J. (2005) 'Recruitment of penicillin-binding protein PBP2 to the division site of *Staphylococcus aureus* is dependent on its transpeptidation substrates', *Molecular Microbiology*, 55(3), pp. 799-807.
- Podmore, A.H. and Reynolds, P.E. (2002) 'Purification and characterization of VanXY(C), a D,D-dipeptidase/D,D-carboxypeptidase in vancomycin-resistant *Enterococcus gallinarum* BM4174', *European Journal of Biochemistry*, 269(11), pp. 2740-6.
- Popieniek, P.H. and Pratt, R.F. (1987) 'A fluorescent ligand for binding studies with glycopeptide antibiotics of the vancomycin class', *Analytical Biochemistry*, 165(1), pp. 108-13.
- Potluri, L., Karczmarek, A., Verheul, J., Piette, A., Wilkin, J.M., Werth, N., Banzhaf, M., Vollmer, W., Young, K.D., Nguyen-Disteche, M. and den Blaauwen, T. (2010) 'Septal and lateral wall localization of PBP5, the major D,D-carboxypeptidase of *Escherichia coli*, requires substrate recognition and membrane attachment', *Molecular Microbiology*, 77(2), pp. 300-23.
- Poxton, I.R., Tarelli, E. and Baddiley, J. (1978) 'The structure of C-polysaccharide from the walls of *Streptococcus pneumoniae*', *Biochemical Journal*, 175(3), pp. 1033-42.
- Pratviel-Sosa, F., Mengin-Lecreux, D. and van Heijenoort, J. (1991) 'Over-production, purification and properties of the uridine diphosphate N-acetylmuramoyl-L-alanine:D-glutamate ligase from *Escherichia coli*', *European Journal of Biochemistry*, 202(3), pp. 1169-76.
- Rasmussen, J.R. and Strominger, J.L. (1978) 'Utilization of a depsipeptide substrate for trapping acyl-enzyme intermediates of penicillin-sensitive D-alanine carboxypeptidases', *Proceedings of the National Academy of Sciences of the United States of America*, 75(1), pp. 84-8.
- Rawlings, N.D., Barrett, A.J. and Bateman, A. (2012) 'MEROPS: the database of proteolytic enzymes, their substrates and inhibitors', *Nucleic Acids Res*, 40(Database issue), pp. D343-50.
- Read, R.J. and Schierbeek, A.J. (1988) 'A phased translation function', *Journal of Applied Crystallography*, 21(5), pp. 490-495.
- Real, G., Fay, A., Eldar, A., Pinto, S.M., Henriques, A.O. and Dworkin, J. (2008) 'Determinants for the subcellular localization and function of a nonessential SEDS protein', *Journal of Bacteriology*, 190(1), pp. 363-76.

- Reith, J. and Mayer, C. (2011) 'Peptidoglycan turnover and recycling in Gram-positive bacteria', *Applied Microbiology and Biotechnology*, 92(1), pp. 1-11.
- Reynolds, P.E., Arias, C.A. and Courvalin, P. (1999) 'Gene vanXYZ encodes D,D -dipeptidase (VanX) and D,D-carboxypeptidase (VanY) activities in vancomycin-resistant *Enterococcus gallinarum* BM4174', *Molecular Microbiology*, 34(2), pp. 341-9.
- Salton, M.R. and Pavlik, J.G. (1960) 'Studies of the bacterial cell wall. VI. Wall composition and sensitivity to lysozyme', *Biochimica et Biophysica Acta*, 39, pp. 398-407.
- Sauvage, E., Derouaux, A., Fraipont, C., Joris, M., Herman, R., Rocaboy, M., Schloesser, M., Dumas, J., Kerff, F., Nguyen-Disteche, M. and Charlier, P. (2014) 'Crystal structure of penicillin-binding protein 3 (PBP3) from *Escherichia coli*', *PLoS One*, 9(5), p. e98042.
- Sauvage, E., Kerff, F., Terrak, M., Ayala, J.A. and Charlier, P. (2008) 'The penicillin-binding proteins: Structure and role in peptidoglycan biosynthesis', *FEMS Microbiology Reviews*, 32(2), pp. 234-258.
- Scapin, G. (2013) 'Molecular replacement then and now', *Acta Crystallogr D Biol Crystallogr*, 69(Pt 11), pp. 2266-75.
- Schaffer, A. and Auld, D.S. (1986) 'Activity of copper-substituted carboxypeptidase A toward oligopeptides and depsipeptides', *Biochemistry*, 25(9), pp. 2476-9.
- Schäffer, C. and Messner, P. (2005) 'The structure of secondary cell wall polymers: how Gram-positive bacteria stick their cell walls together', *Microbiology*, 151(3), pp. 643-651.
- Scheffers, D.J., Jones, L.J. and Errington, J. (2004) 'Several distinct localization patterns for penicillin-binding proteins in *Bacillus subtilis*', *Molecular Microbiology*, 51(3), pp. 749-64.
- Scheffers, D.J. and Pinho, M.G. (2005) 'Bacterial cell wall synthesis: new insights from localization studies', *Microbiology and Molecular Biology Reviews*, 69(4), pp. 585-607.
- Schirner, K. and Errington, J. (2009) 'Influence of heterologous MreB proteins on cell morphology of *Bacillus subtilis*', *Microbiology*, 155(Pt 11), pp. 3611-21.
- Schneewind, O. and Missiakas, D.M. (2012) 'Protein secretion and surface display in Gram-positive bacteria', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1592), pp. 1123-1139.
- Schneider, T., Senn, M.M., Berger-Bachi, B., Tossi, A., Sahl, H.G. and Wiedemann, I. (2004) 'In vitro assembly of a complete, pentaglycine interpeptide bridge containing cell wall precursor (lipid II-Gly5) of *Staphylococcus aureus*', *Molecular Microbiology*, 53(2), pp. 675-85.
- Schrodinger, LLC (2010) *The PyMOL Molecular Graphics System, Version 1.4.1*.
- Schuster, C., Dobrinski, B. and Hakenbeck, R. (1990) 'Unusual septum formation in *Streptococcus pneumoniae* mutants with an alteration in the D,D-carboxypeptidase penicillin-binding protein 3', *Journal of Bacteriology*, 172(11), pp. 6499-505.

- Sham, L.T., Butler, E.K., Lebar, M.D., Kahne, D., Bernhardt, T.G. and Ruiz, N. (2014) 'Bacterial cell wall. MurJ is the flippase of lipid-linked precursors for peptidoglycan biogenesis', *Science*, 345(6193), pp. 220-2.
- Shannon, R. (1976) 'Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides', *Acta Crystallographica Section A*, 32(5), pp. 751-767.
- Sheldrick, G.M. (2007) 'A short history of SHELX', *Acta Crystallographica Section A: Foundations of Crystallography*, 64(1), pp. 112-122.
- Sheldrick, G.M. (2008) 'A short history of SHELX', *Acta Crystallogr A*, 64(Pt 1), pp. 112-22.
- Shockman GD, Barrett JF (1983) Structure, function, and assembly of cell walls of gram-positive bacteria. *Annual Review of Microbiology* 37, pp. 501-527
- Sieger, B., Schubert, K., Donovan, C. and Bramkamp, M. (2013) 'The lipid II flippase RodA determines morphology and growth in *Corynebacterium glutamicum*', *Molecular Microbiology*, 90(5), pp. 966-82.
- Simanshu, D.K., Savithri, H.S. and Murthy, M.R. (2005) 'Crystal structures of ADP and AMPPNP-bound propionate kinase (TdcD) from *Salmonella typhimurium*: comparison with members of acetate and sugar kinase/heat shock cognate 70/actin superfamily', *Journal of Molecular Biology*, 352(4), pp. 876-92.
- Smith, A.A., Carlow, D.C., Wolfenden, R. and Short, S.A. (1994) 'Mutations affecting transition-state stabilization by residues coordinating zinc at the active site of cytidine deaminase', *Biochemistry*, 33(21), pp. 6468-74.
- Snowden, M.A. and Perkins, H.R. (1990) 'Peptidoglycan cross-linking in *Staphylococcus aureus*. An apparent random polymerisation process', *European Journal of Biochemistry*, 191(2), pp. 373-7.
- Spratt, B.G. (1975) 'Distinct penicillin binding proteins involved in the division, elongation, and shape of *Escherichia coli* K12', *Proceedings of the National Academy of Sciences of the United States of America*, 72(8), pp. 2999-3003.
- Spratt, B.G., Zhou, J., Taylor, M. and Merrick, M.J. (1996) 'Monofunctional biosynthetic peptidoglycan transglycosylases', *Molecular Microbiology*, 19(3), pp. 639-40.
- Stahlberg, J., Divne, C., Koivula, A., Piens, K., Claeysens, M., Teeri, T.T. and Jones, T.A. (1996) 'Activity studies and crystal structures of catalytically deficient mutants of cellobiohydrolase I from *Trichoderma reesei*', *J Mol Biol*, 264(2), pp. 337-49.
- Stamatakis, A. (2006) 'RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models', *Bioinformatics*, 22(21), pp. 2688-90.
- Stamatakis, A., Hoover, P. and Rougemont, J. (2008) 'A rapid bootstrap algorithm for the RAxML Web servers', *Syst Biol*, 57(5), pp. 758-71.
- Steitz, T.A., Ludwig, M.L., Quijcho, F.A. and Lipscomb, W.N. (1967) 'The structure of carboxypeptidase A. V. Studies of enzyme-substrate and enzyme-inhibitor complexes at 6 Å resolution', *Journal of Biological Chemistry*, 242(20), pp. 4662-8.

- Stierand, K., Maass, P.C. and Rarey, M. (2006) 'Molecular complexes at a glance: automated generation of two-dimensional complex diagrams', *Bioinformatics*, 22(14), pp. 1710-6.
- Stierand, K. and Rarey, M. (2010) 'Drawing the PDB: Protein-Ligand Complexes in Two Dimensions', *ACS Med Chem Lett*, 1(9), pp. 540-5.
- Strominger, J.L. (1958) 'Enzymic transfer of pyruvate to uridine diphosphoacetylglucosamine', *Biochimica et Biophysica Acta*, 30(3), pp. 645-6.
- Sung, M.T., Lai, Y.T., Huang, C.Y., Chou, L.Y., Shih, H.W., Cheng, W.C., Wong, C.H. and Ma, C. (2009) 'Crystal structure of the membrane-bound bifunctional transglycosylase PBP1b from *Escherichia coli*', *Proceedings of the National Academy of Sciences of the United States of America*, 106(22), pp. 8824-9.
- Sycuro, L.K., Rule, C.S., Petersen, T.W., Wyckoff, T.J., Sessler, T., Nagarkar, D.B., Khalid, F., Pincus, Z., Biboy, J., Vollmer, W. and Salama, N.R. (2013) 'Flow cytometry-based enrichment for cell shape mutants identifies multiple genes that influence *Helicobacter pylori* morphology', *Molecular Microbiology*, 90(4), pp. 869-883.
- Tamas, M.J., Sharma, S.K., Ibstedt, S., Jacobson, T. and Christen, P. (2014) 'Heavy metals and metalloids as a cause for protein misfolding and aggregation', *Biomolecules*, 4(1), pp. 252-67.
- Tatar, L.D., Marolda, C.L., Polischuk, A.N., van Leeuwen, D. and Valvano, M.A. (2007) 'An *Escherichia coli* undecaprenyl-pyrophosphate phosphatase implicated in undecaprenyl phosphate recycling', *Microbiology*, 153(Pt 8), pp. 2518-29.
- Templin, M.F., Ursinus, A. and Holtje, J.V. (1999) 'A defect in cell wall recycling triggers autolysis during the stationary growth phase of *Escherichia coli*', *EMBO Journal*, 18(15), pp. 4108-17.
- Ten Eyck, L.F. (1985) 'Fast Fourier transform calculation of electron density maps', *Methods in Enzymology*, 115, pp. 324-37.
- Turkenburg, J.P. and Dodson, E.J. (1996) 'Modern developments in molecular replacement', *Curr Opin Struct Biol*, 6(5), pp. 604-10.
- Turski, M.L., Brady, D.C., Kim, H.J., Kim, B.E., Nose, Y., Counter, C.M., Winge, D.R. and Thiele, D.J. (2012) 'A novel role for copper in Ras/mitogen-activated protein kinase signaling', *Molecular and Cellular Biology*, 32(7), pp. 1284-95.
- Umbreit, J.N. and Strominger, J.L. (1972) 'Complex lipid requirements for detergent-solubilized phosphoacetylmuramyl-pentapeptide translocase from *Micrococcus luteus*', *Proceedings of the National Academy of Sciences of the United States of America*, 69(7), pp. 1972-4.
- Uversky, V.N., Li, J. and Fink, A.L. (2001) 'Metal-triggered structural transformations, aggregation, and fibrillation of human alpha-synuclein. A possible molecular link between Parkinson's disease and heavy metal exposure', *Journal of Biological Chemistry*, 276(47), pp. 44284-96.

- Veiga, P., Piquet, S., Maisons, A., Furlan, S., Courtin, P., Chapot-Chartier, M.P. and Kulakauskas, S. (2006) 'Identification of an essential gene responsible for D-Asp incorporation in the *Lactococcus lactis* peptidoglycan crossbridge', *Molecular Microbiology*, 62(6), pp. 1713-24.
- Vinella, D., Joseleau-Petit, D., Thevenet, D., Boulloc, P. and D'Ari, R. (1993) 'Penicillin-binding protein 2 inactivation in *Escherichia coli* results in cell division inhibition, which is relieved by FtsZ overexpression', *Journal of Bacteriology*, 175(20), pp. 6704-10.
- Vollmer, W., Blanot, D. and De Pedro, M.A. (2008a) 'Peptidoglycan structure and architecture', *FEMS Microbiology Reviews*, 32(2), pp. 149-167.
- Vollmer, W., Joris, B., Charlier, P. and Foster, S. (2008b) 'Bacterial peptidoglycan (murein) hydrolases', *FEMS Microbiology Reviews*, 32(2), pp. 259-286.
- Votsch, W. and Templin, M.F. (2000) 'Characterization of a beta -N-acetylglucosaminidase of *Escherichia coli* and elucidation of its role in muropeptide recycling and beta -lactamase induction', *Journal of Biological Chemistry*, 275(50), pp. 39032-8.
- Waldron, K.J. and Robinson, N.J. (2009) 'How do bacterial cells ensure that metalloproteins get the correct metal?', *Nat Rev Microbiol*, 7(1), pp. 25-35.
- Wang, L., Khattar, M.K., Donachie, W.D. and Lutkenhaus, J. (1998) 'FtsI and FtsW are localized to the septum in *Escherichia coli*', *Journal of Bacteriology*, 180(11), pp. 2810-6.
- Wang, X., Yang, X., Yang, C., Wu, Z., Xu, H. and Shen, Y. (2011) 'Crystal structure of outer membrane protein NMB0315 from *Neisseria meningitidis*', *PLoS One*, 6(10), p. e26845.
- Wei, Y., Havasy, T., McPherson, D.C. and Popham, D.L. (2003) 'Rod shape determination by the *Bacillus subtilis* class B penicillin-binding proteins encoded by pbpA and pbpH', *Journal of Bacteriology*, 185(16), pp. 4717-26.
- Weiss, M. (2001) 'Global indicators of X-ray data quality', *Journal of Applied Crystallography*, 34(2), pp. 130-135.
- Wilkins, M.R., Gasteiger, E., Bairoch, A., Sanchez, J.C., Williams, K.L., Appel, R.D. and Hochstrasser, D.F. (1999) 'Protein identification and analysis tools in the ExPASy server', *Methods Mol Biol*, 112, pp. 531-52.
- Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A. and Wilson, K.S. (2011) 'Overview of the CCP4 suite and current developments', *Acta Crystallographica Section D: Biological Crystallography*, 67(4), pp. 235-242.
- Wood, W.B. (1966) 'Host specificity of DNA produced by *Escherichia coli*: bacterial mutations affecting the restriction and modification of DNA', *J Mol Biol*, 16(1), pp. 118-133.
- Wright, G.D., Molinas, C., Arthur, M., Courvalin, P. and Walsh, C.T. (1992) 'Characterization of VanY, a DD-carboxypeptidase from vancomycin-resistant *Enterococcus faecium* BM4147', *Antimicrobial Agents and Chemotherapy*, 36(7), pp. 1514-1518.
- Wu, S., Zhang, C., Xu, D. and Guo, H. (2010) 'Catalysis of carboxypeptidase A: promoted-water versus nucleophilic pathways', *J Phys Chem B*, 114(28), pp. 9259-67.

- Xu, X., Zhang, L., Shen, D., Wu, H., Peng, L. and Li, J. (2009) 'Effect of metal ion substitutions in anticoagulation factor I from the venom of *Agkistrodon acutus* on the binding of activated coagulation factor X and on structural stability', *J Biol Inorg Chem*, 14(4), pp. 559-71.
- Yeats, C., Finn, R.D. and Bateman, A. (2002) 'The PASTA domain: a beta-lactam-binding domain', *Trends in Biochemical Sciences*, 27(9), p. 438.
- Yim, G., Thaker, M.N., Koteva, K. and Wright, G. (2014) 'Glycopeptide antibiotic biosynthesis', *Journal of Antibiotics*, 67(1), pp. 31-41.
- Yocum, R.R., Rasmussen, J.R. and Strominger, J.L. (1980) 'The mechanism of action of penicillin. Penicillin acylates the active site of *Bacillus stearothermophilus* D-alanine carboxypeptidase', *Journal of Biological Chemistry*, 255(9), pp. 3977-86.
- Zhou, Z.S., Peariso, K., Penner-Hahn, J.E. and Matthews, R.G. (1999) 'Identification of the zinc ligands in cobalamin-independent methionine synthase (MetE) from *Escherichia coli*', *Biochemistry*, 38(48), pp. 15915-26.

Appendix

Appendix 1: General reagents, buffers and growth media

Electrophoresis:

- 1 x TAE: 0.4 M Tris-acetate (pH 8.0), 0.01 M EDTA.
- 6 x DNA loading dye: 0.25 % (w/v) bromophenol blue, 0.25 % (w/v) xylene cyanol FF, 40 % (w/v) sucrose.
- 1 x SDS running buffer: 40 mM Tris-HCl (pH 8.5), 250 mM glycine, 0.5 % (w/v) SDS.
- 6 x SDS loading dye: 50 mM Tris-HCl (pH 6.8), 2 % (w/v) SDS, 0.1 % (w/v) bromophenol blue, 10 % (v/v) glycerol, 1 mM DTT.
- 1 x native running buffer: 40 mM Tris-HCl (pH 8.5), 250 mM glycine.
- 5 x native loading dye: 40 mM Tris-HCl (pH 6.8), 0.06 % (w/v) bromophenol blue, 13 % (v/v) glycerol.
- Coomassie blue stain: 2.5 g Coomassie blue R-250 in 1 l of destain.
- Destain: 20 % (v/v) ethanol, 10 % (v/v) glacial acetic acid in 1 l H₂O.

Stacking Gel (5 %)

5 % (v/v) Bis-acrylamide, 0.2 M Tris (pH 6.8), 0.1 % (v/v) sodium dodecyl sulphate (SDS), 0.1 % (v/v) ammonium persulphate, 0.1 % (v/v) Tetramethylethylenediamine (TEMED)

Resolving Gel (15 %)

15 % (v/v) Bis-acrylamide, 0.35 M Tris (pH 8.8), 0.1 % (v/v) SDS, 0.1 % (v/v) ammonium persulphate, 0.1 % (v/v) TEMED

Luria Bertani broth (LB): 1 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 0.5 % (w/v) NaCl.

LB Agar plates: As LB broth, but with the addition of 1.5 % (w/v) agar prior to autoclaving.

Selenomethionine growth media:

20 x M9 media: 0.4 M NH₄Cl, 0.4 M KH₂PO₄, 0.85 M Na₂HPO₄.

Amino acids mix I: Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Pro, Ser, Thr, Val (all at 0.4 g/ml)

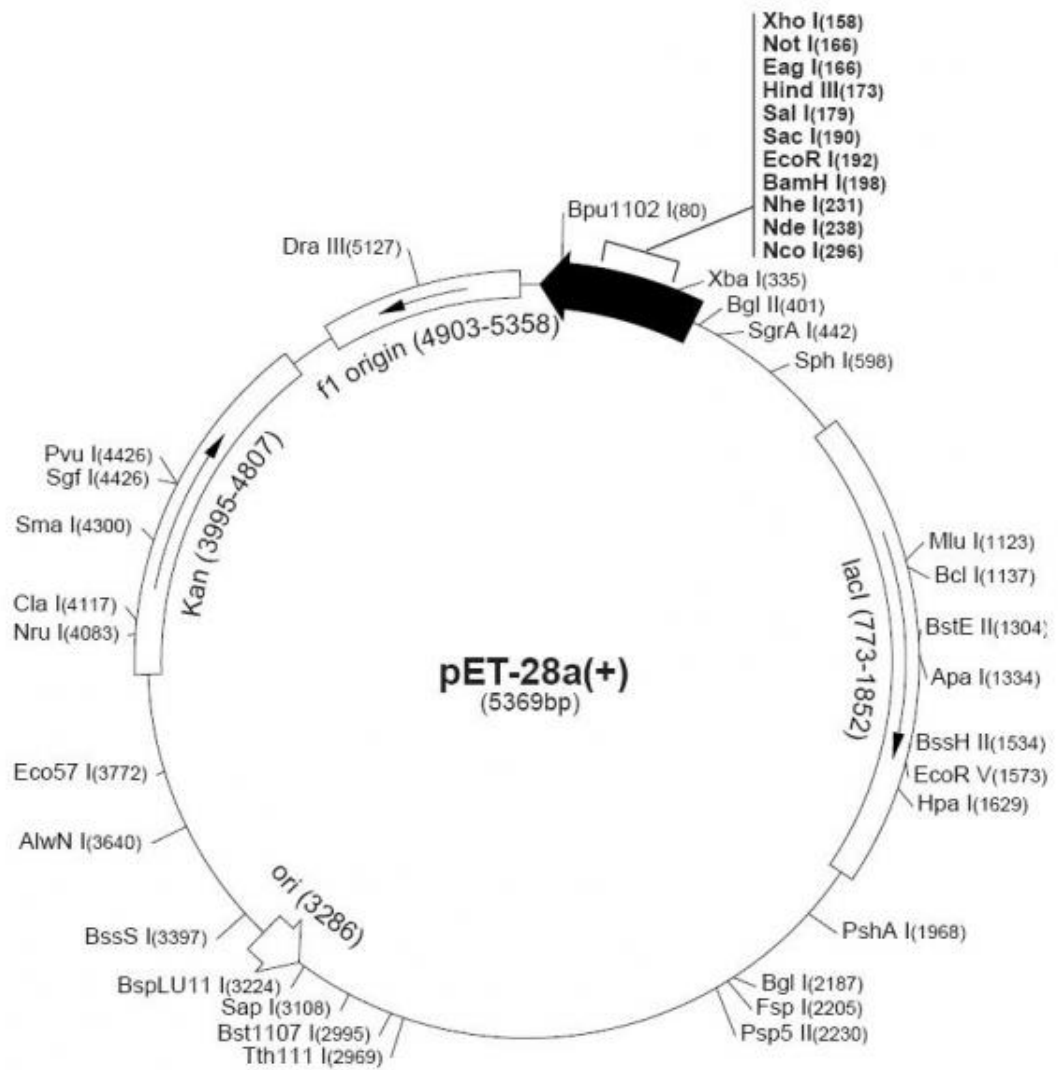
Amino acids mix II: Phe, Trp, Tyr, Asp (all at 0.4 g/ml)

Vitamin solution: niacinamide, pyroxidine monohydrochloride, riboflavin, thiamine

(all at 1 mg/ml).

Se-met media: 2 mM MgSO₄, 2 x M9 media, 45 uM FeSO₄, 0.4 % (w/v) glucose, 0.1 % (v/v) amino acid mix I, 0.1 % (v/v) amino acid mix II, 0.01 % (v/v) vitamin solution, Se-L-Met (0.4 g/l).

Appendix 2: Plasmid map of pET28a



Appendix Figure 2.1: Plasmid map of pET28a. The coding sequences for the various proteins were inserted into the multiple cloning site, shown as a black arrow. Image taken from the Merck Millipore website http://www.merckmillipore.de/life-science-research/vector-table-novagen-pet-vector-table/c_HdSb.s1O77QAAAEhPqsLdcab

Appendix 3: Primers used in this study

Appendix Table 3.1: List of mutagenesis primers used in this study. Mutations are underlined, with specific base changes in bold.

Primer Name	Sequence	Mutation Introduced
R120AFw	GCGATCATTACAGTGGTTTT <u>GCA</u> AGTTATGAAACTCAGACCAAATC	Arg to Ala substitution
R120ARv	GAGTTTGGTCTGAGTTTCATAACT <u>TGC</u> AAAACCACTGTAATGATCGC	Arg to Ala substitution
E204AFw	GAAACAGGCTATATGGCT <u>GCG</u> GAATGGCACCTGCGTTATG	Glu to Ala substitution
E204ARv	CATAACGCAGGTGCCATTC <u>CGC</u> AGCCATATAGCCTGTTTC	Glu to Ala substitution
E204HFw	GAAACAGGCTATATGGCT <u>CAC</u> GAATGGCACCTGCGTTATG	Glu to His substitution
E204HRv	CATAACGCAGGTGCCATTC <u>GTG</u> AGCCATATAGCCTGTTTC	Glu to His substitution

Appendix 4: Co-crystallisation conditions of *SpLdcB* with a range of ligands

Appendix table 4.1: Crystallisation conditions for *SpLdcB* with a range of ligands.

Ligand	Condition	Components
M-Tri-Lys	PACT C8	0.2 M ammonium chloride, 0.1 M HEPES pH 7, 20% w/v PEG 6000
	MORPH D8	0.12 M alcohols (1,6-hexanediol; 1-butanol; 1,2-propanediol (racemic); 2-propanol; 1,4-butanediol; 1,3-propanediol), 0.1 M buffer system 2 pH 7.5 (sodium HEPES; MOPS (acid)), 37.50% v/v MPD_PEG1000_PEG3350
	INDEX G9	0.1 M potassium thiocyanate, 30 % w/v PEG 2000 MME
	MORPH B8	0.09 M halogens (NaF; NaBr; NaI), 0.1 M buffer system 2 pH 7.5 (sodium HEPES; MOPS (acid)), 37.50% v/v MPD_PEG1000_PEG3350
M-Tri-Lys(D-Asn)	MORPH D8	0.12 M alcohols (1,6-hexanediol; 1-butanol; 1,2-propanediol (racemic); 2-propanol; 1,4-butanediol; 1,3-propanediol), 0.1 M buffer system 2 pH 7.5 (sodium HEPES; MOPS (acid)), 37.50% v/v MPD_PEG1000_PEG3350
	PACT C8	0.2 M ammonium chloride, 0.1 M HEPES pH 7, 20% w/v PEG 6000
	PACT D8	0.2 M ammonium chloride, 0.1 M tris pH 8.0, 20 % w/v PEG 6000
	MORPH E8	0.12 M ethylene glycols (di-ethyleneglycol; tri-ethyleneglycol; tetra-ethyleneglycol; penta-ethyleneglycol), 0.1 M buffer system 2 pH 7.5 (sodium HEPES; MOPS (acid)), 37.50% v/v MPD_PEG1000_PEG3350
Tri-DAP	PACT C8	0.2 M ammonium chloride, 0.1 M HEPES pH 7, 20% PEG 6000
	STRUCTURE G4	0.1 M Na HEPES pH 7.5, 20 % w/v PEG 10,000
	MORPH F5	0.12 M monosaccharides (D-glucose; D-mannose; D-galactose; L-fucose; D-xylose; N-acetyl-D-glucosamine), 0.1 M buffer system 2 pH 7.5 (sodium HEPES; MOPS (acid)), 30.00% v/v PEG500MME_PEG20,000
MDP	PACT C8	0.2 M ammonium chloride, 0.1 M HEPES pH 7, 20% PEG 6000
	STRUCTURE G8	0.1 M Na citrate pH 5.6, 35 % v/v tert-butanol
	MORPH E8	0.12 M ethylene glycols (di-ethyleneglycol; tri-ethyleneglycol; tetra-ethyleneglycol; penta-ethyleneglycol), 0.1 M buffer system 2 pH 7.5 (sodium HEPES; MOPS (acid)), 37.50% v/v MPD_PEG1000_PEG3350

Structure of the LdcB LD-Carboxypeptidase Reveals the Molecular Basis of Peptidoglycan Recognition

Christopher N. Hoyland,¹ Christine Aldridge,² Robert M. Cleverley,¹ Marie-Clémence Duchêne,³ George Minasov,^{4,5} Olena Onopriyenko,^{5,6} Karzan Sidiq,² Peter J. Stogios,^{5,6} Wayne F. Anderson,^{4,5} Richard A. Daniel,² Alexei Savchenko,^{5,6} Waldemar Vollmer,² and Richard J. Lewis^{1,*}

¹Institute for Cell and Molecular Biosciences, Newcastle University, Newcastle upon Tyne NE2 4HH, UK

²Centre for Bacterial Cell Biology, Institute for Cell and Molecular Biosciences, Newcastle University, Newcastle upon Tyne NE2 4AX, UK

³Institut des Sciences de la Vie, Université Catholique de Louvain, 1348 Louvain-la-Neuve, Belgium

⁴Department of Molecular Pharmacology and Biological Chemistry, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA

⁵Center for Structural Genomics of Infectious Diseases (CSGID)

⁶Department of Chemical Engineering and Applied Chemistry, 200 College Street, University of Toronto, Toronto, ON M5G 1L6, Canada

*Correspondence: r.lewis@ncl.ac.uk

<http://dx.doi.org/10.1016/j.str.2014.04.015>

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/3.0/>).

SUMMARY

Peptidoglycan surrounds the bacterial cytoplasmic membrane to protect the cell against osmolysis. The biosynthesis of peptidoglycan, made of glycan strands crosslinked by short peptides, is the target of antibiotics like β -lactams and glycopeptides. Nascent peptidoglycan contains pentapeptides that are trimmed by carboxypeptidases to tetra- and tripeptides. The well-characterized DD-carboxypeptidases hydrolyze the terminal D-alanine from the stem pentapeptide to produce a tetrapeptide. However, few LD-carboxypeptidases that produce tripeptides have been identified, and nothing is known about substrate specificity in these enzymes. We report biochemical properties and crystal structures of the LD-carboxypeptidases LdcB from *Streptococcus pneumoniae*, *Bacillus anthracis*, and *Bacillus subtilis*. The enzymes are active against bacterial cell wall tetrapeptides and adopt a zinc-carboxypeptidase fold characteristic of the LAS superfamily. We have also solved the structure of *S. pneumoniae* LdcB with a product mimic, elucidating the residues essential for peptidoglycan recognition and the conformational changes that occur on ligand binding.

INTRODUCTION

Bacteria surround their cytoplasmic membrane with a large, semipermeable barrier, the peptidoglycan (PG) sacculus, which not only protects the cell from lysis from turgor but also helps to maintain cell shape. PG comprises repeating units of the *N*-acetylglucosamine-*N*-acetylmuramic acid (GlcNAc—MurNAc) disaccharide that polymerizes into long glycan chains. Pentapeptides, usually comprising L-Ala—D- γ -Glu[Gln]—L-Lys—D-Ala—

D-Ala, extend from the MurNAc moiety. In Gram-negative bacteria and some Gram-positive Firmicutes, the L-Lys at position 3 is replaced by *meso*-diaminopimelic acid (A₂pm), which is amidated in *Bacillus subtilis* (Vollmer et al., 2008a). Peptides protruding from adjacent glycan chains are crosslinked by transpeptidases to form the net-like PG sacculus, which the cell has to enlarge in order to grow and to divide (Sauvage et al., 2008). Newly inserted PG matures and is turned over by PG hydrolases, and the balance between synthesis and degradation of PG is key to normal cell division (Vollmer et al., 2008b). The importance of PG to the healthy physiology of a bacterium is highlighted by the fact that its biosynthesis is the target of some of our most successful antibiotics, β -lactams such as penicillin and glycopeptides such as vancomycin, resistance to which in pathogens is an urgent and currently unmet human health challenge.

In most species, the terminal D-alanine of the nascent pentapeptide stem is hydrolysed by DD-carboxypeptidases from the class C penicillin binding protein (PBP) family (Lovering et al., 2012). LD-carboxypeptidases subsequently hydrolyze the L-Lys[A₂pm]—D-Ala bond to release the tripeptide. Though tripeptides are found in many bacteria, very few of the corresponding LD-carboxypeptidases have been identified. Membrane preparations of *B. subtilis* have shown strong LD-carboxypeptidase activity against cell-wall-derived tetrapeptides (Arminjon et al., 1977), though the specific enzyme involved has not been isolated. The best characterized LD-carboxypeptidase is LD-carboxypeptidase A (LdcA), a cytoplasmic, penicillin-insensitive member of the S66 family of serine peptidases that is involved in PG recycling and is essential for cells entering the stationary phase (Metz et al., 1986; Templin et al., 1999). The cytoplasmic localization of LdcA precludes its involvement in the production of cell wall tripeptides. Other LD-carboxypeptidases either must be active within the periplasm of Gram-negative bacteria or are anchored to the exterior of the cell membrane in Gram-positives.

Several LD-carboxypeptidase genes have been identified recently, including *dacB* from *Lactococcus lactis* and *Streptococcus pneumoniae* (Barendt et al., 2011; Courtin et al., 2006), *csd6* from *Helicobacter pylori* (Sycuro et al., 2012) and *pgp2*

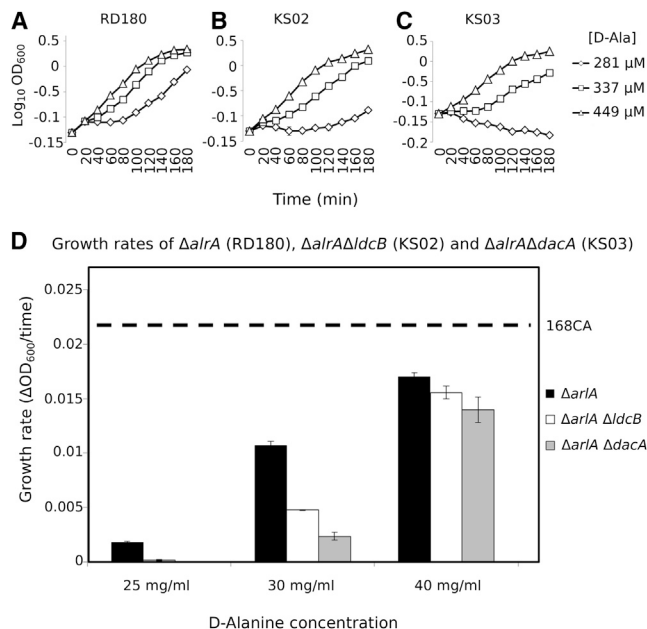


Figure 1. Effect of *dacA* and *ldcB* Alleles in a D-alanine Auxotroph
(A–C) The growth of strains (A) RD180 ($\Delta alrA$), (B) KS02 ($\Delta ldcB \Delta alrA$), and (C) KS03 ($\Delta dacA \Delta alrA$) incubated at 37°C in LB medium supplemented with different concentrations of D-alanine was determined by monitoring the culture at OD_{600} and plotting it against time of incubation.
(D) The calculated growth rate of each strain at three different concentrations of exogenous D-alanine. The values shown represent the average of three independent experiments. The dashed line shows the growth rate of wild-type *B. subtilis* strain 168CA under the same conditions. The error bars represent the SD of growth rate means at a confidence level of 95%.

from *Campylobacter jejuni* (Firdich et al., 2014), though it should be noted that *csd6* and *pgp2* share no sequence similarity with *dacB*. The corresponding null mutants exhibited increased amounts of tetrapeptides in the cell wall and a concomitant decrease in tripeptide content (Sycuro et al., 2012; Firdich et al., 2014). The *S. pneumoniae* *dacB* mutant also displayed severe morphological defects, developing into small, spherical cells or elongating into long, thick cells, as a result of division asymmetry produced by random septum localization (Barendt et al., 2011). Both the *H. pylori* *csd6* and the *C. jejuni* *pgp2* mutants have a morphological defect, displaying a straight shape instead of the normal helical morphology, indicating that *csd6* and *pgp2* also contribute to cell shape maintenance (Firdich et al., 2014; Sycuro et al., 2012). In *C. jejuni* (Firdich et al., 2012) and *H. pylori* (Sycuro et al., 2010), cell shape is also linked to pathogenicity, and PG fragments released by growing cells are important signaling molecules for immune recognition and modulation in animals ranging from insects to man (Royet et al., 2011). Therefore, the PG plays an important role in bacterial shape maintenance, viability, virulence and immune signaling.

In this study, we have investigated the biochemical activity of *S. pneumoniae* and *B. subtilis* DacB against cell wall sacculi and synthetic tetrapeptides. Both proteins are shown to be active against cell-wall-derived tetrapeptides and synthetic tetrapeptides lacking the sugar moiety but are inactive against tetra-

peptides terminating in L-alanine, confirming their classification as LD-carboxypeptidases. We therefore propose to rename these enzymes LD-carboxypeptidase B (LdcB). We present the crystal structures of both LdcB enzymes and the ortholog from *Bacillus anthracis*, all of which belong to the LAS (lysostaphin, D-Ala-D-Ala metallopeptidases, sonic hedgehog) family of zinc-dependent peptidases. The structure of *S. pneumoniae* LdcB with MurNAc–L-Ala–D- γ -Gln–L-Lys–(D-Asn) bound reveals the molecular basis of PG recognition and the key roles played in catalysis by active site elements. It is surprising that the active site is reminiscent, yet reversed, of that found in pancreatic carboxypeptidase A, suggesting that the enzymes have converged on a common catalytic mechanism.

RESULTS

Role for LdcB In Vivo

The mature PG of *S. pneumoniae* and *B. subtilis* contains mostly tripeptides, indicating the presence of PG-active peptidases in these organisms (Atrih et al., 1999; Bui et al., 2012). Both species contain DacA orthologs, DD-carboxypeptidases that cleave pentapeptides to tetrapeptides; however, the LD-carboxypeptidases that produce tripeptides from the tetrapeptides have remained elusive until recently. The PG of a *S. pneumoniae* *dacB* mutant contained mainly tetrapeptides, suggesting strongly that SpDacB is an LD-carboxypeptidase (hereinafter renamed as SpLdcB) (Barendt et al., 2011). Single mutants of the *dacA* and *dacB* orthologs in *B. subtilis*, *dacA*, and *yodJ* (hereinafter renamed as *ldcB*), had no obvious phenotype during vegetative growth, indicating that the trimming of peptides or the released D-Ala is not required for growth under laboratory conditions (Atrih et al., 1999). To test whether the released D-Ala becomes important under D-Ala limiting conditions, we constructed *dacA* and *ldcB* null mutants in a *B. subtilis* strain auxotrophic for D-Ala (an alanine racemase [*alrA*] null mutant) and assayed growth in the presence of D-Ala. The D-Ala-dependent strain, RD180, required at least 450 μ M D-Ala for normal growth, and it grew more slowly at lower D-Ala concentrations (Figure 1). However, strains also bearing either the *dacA* (KS03) or the *ldcB* (KS02) alleles exhibited even more severe growth rate dependencies on D-Ala, with the *dacA* mutant the most severely affected. These data suggest strongly that LdcB is an LD-carboxypeptidase that releases D-Ala from the cell wall, which can be used to support growth under D-Ala-limiting conditions.

To test directly whether LdcB functions as an LD-carboxypeptidase, we generated *S. pneumoniae* and *B. subtilis* *ldcB* mutants and isolated their cell wall PG. Muropeptides were released with cellosyl and analyzed by HPLC (Figures 2A and 2B). The *ldcB* mutant strains lacked tripeptides and contained mainly tetrapeptides in their PGs, consistent with previous work indicating that LdcB is the main enzyme trimming cell wall tetrapeptides (Barendt et al., 2011). To obtain more direct evidence for an enzymatic activity, purified, recombinant LdcB enzymes from *S. pneumoniae* and *B. subtilis* (termed SpLdcB and BsLdcB) were incubated with PG isolated from the respective *ldcB* mutant strains before digesting with cellosyl and analyzing the resulting muropeptides by HPLC (Figures 2A and 2B). Both SpLdcB and BsLdcB almost quantitatively digested PG tetrapeptides with a

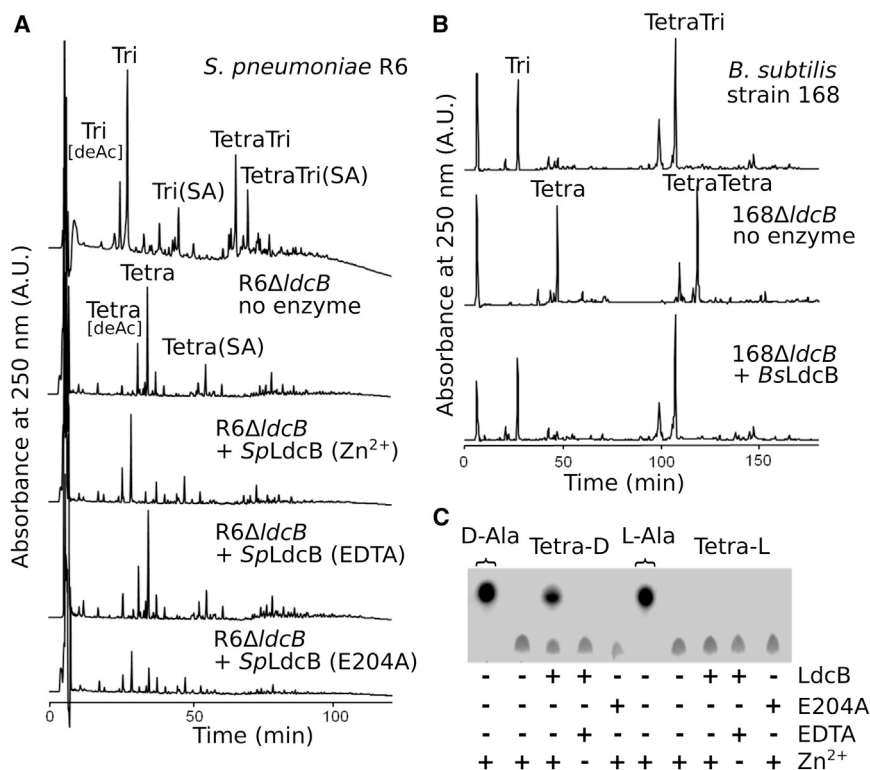


Figure 2. Activity of *SpLdcB* and *BsLdcB* against Peptidoglycan and Peptides

(A) HPLC chromatograms of mucopeptides from *S. pneumoniae* strains, obtained after incubating PG with or without *SpLdcB* or *SpLdcB*(E204A) in the presence of Zn²⁺ or EDTA, followed by digestion with cellosyl and reduction with sodium borohydride. The mucopeptides are: Tri, GlcNAc–MurNAc(r)–L-Ala–D-γ-Gln–L-Lys; Tetra, GlcNAc–MurNAc(r)–L-Ala–D-γ-Gln–L-Lys–D-Ala; TetraTri, GlcNAc–MurNAc–L-Ala–D-γ-Gln–L-Lys–D-Ala–L-Lys–D-γ-Gln–L-Ala–MurNAc(r)–GlcNAc; TetraTetra, GlcNAc–MurNAc–L-Ala–D-γ-Gln–L-Lys–(D-Ala)–D-Ala–L-Lys–D-γ-Gln–L-Ala–MurNAc(r)–GlcNAc; deAc, deacetylation at GlcNAc; SA, L-Ser–L-Ala branch (at L-Lys). GlcNAc, *N*-acetylglucosamine; MurNAc(r), *N*-acetylmuramitol.

(B) HPLC chromatograms of mucopeptides from *B. subtilis* strains, obtained after incubating PG with or without *BsLdcB*, followed by digestion with cellosyl and reduction with sodium borohydride. Mucopeptides: Tri, GlcNAc–MurNAc(r)–L-Ala–D-γ-Glu–*meso*-Dap(NH₂); Tetra, GlcNAc–MurNAc(r)–L-Ala–D-γ-Glu–*meso*-Dap(NH₂)–D-Ala; TetraTri, GlcNAc–MurNAc–L-Ala–D-γ-Glu–*meso*-Dap(NH₂)–D-Ala–*meso*-Dap(NH₂)–D-γ-Glu–L-Ala–MurNAc(r)–GlcNAc; TetraTetra, GlcNAc–MurNAc–L-Ala–D-γ-Glu–*meso*-Dap(NH₂)–(D-Ala)–D-Ala–*meso*-Dap(NH₂)–D-γ-Glu–L-Ala–MurNAc(r)–GlcNAc; *meso*-Dap(NH₂), *meso*-diaminopimelic acid (amidated).

(C) TLC of *SpLdcB* incubated with various substrates.

carboxy-terminal D-Ala producing the tripeptides, proving their LD-carboxypeptidase activity. *SpLdcB* was active in the presence of Zn²⁺ ions and inactive in the presence of EDTA, suggesting a metal-ion-dependent peptidase mechanism.

SpLdcB and BsLdcB Are Catalytically Active against Synthetic Cell Wall Tetrapeptides

In addition to removing the terminal D-Ala from *S. pneumoniae* cell wall PG tetrapeptides (Figure 2A), *SpLdcB* was also tested against synthetic tetrapeptides (L-Ala–D-Gln–L-Lys–D-Ala [tetra-D]) and its L-isomer, L-Ala–D-Gln–L-Lys–L-Ala [tetra-L]). Though these substrates lack the glycan component and differ further from the natural substrate by the replacement of D-γ-glutamine with D-glutamine, they have the potential to discriminate between LD- and LL-carboxypeptidase activities. *SpLdcB* was active solely against the tetra-D ligand (Figure 2C), an activity that was dependent on the presence of zinc and abrogated totally in the presence of EDTA (Figure 2C).

Overall Structure of LdcB

The crystal structures of *SpLdcB* and *Bacillus anthracis* LdcB (*BaLdcB*) were solved by single wavelength anomalous dispersion (SAD); *BsLdcB* was solved by molecular replacement. The LdcB enzymes are single domain proteins comprising two subdomains separated by a V-shaped cleft (Figures 3A–3C). The first subdomain comprises a four-stranded, antiparallel β sheet (β4–β7) that opposes the cleft, and the outside face of the β sheet is flanked by three α helices (α1, α4, and α6). The second subdo-

main comprises an N-terminal, three-stranded, antiparallel β sheet (β1–β3), somewhat discrete from the four helical bundle (α2, α3, α5, and α7) that makes up the rest of the subdomain. Though the *BaLdcB* structure contains the same subdomain, it is less well structured in this region than *BsLdcB* and *SpLdcB*, and, as a consequence, the β1–β3 sheet is not formed. Neither subdomain is formed by a contiguous, linear part of the primary sequence; six loops (β3–α1, β4–α2, α3–β5, β6–α5, α5–β7, and α6–α7) connect the two subdomains.

The cleft between subdomains is 15–18 Å wide at the top and 15 Å deep and narrows down to the base where a Zn²⁺ ion, added to the enzyme during purification, is located (Figures 3A–3C), which is coordinated in a tetrahedral manner by three amino acids. In *SpLdcB*, the liganding residues are His153, Asp160, and His207 (Figure 3D); the equivalents in *BsLdcB*/*BaLdcB* are His185/His160, Asp192/Asp167, and His244/His219 (Figures 3E and 3F). These zinc ligands are universally conserved in LdcB sequences, consistent with an important functional role in these enzymes. The active site of *SpLdcB* contains unattributed residual electron density that occupies the position of the fourth ligand in the tetrahedral coordination sphere of the zinc (Figure 3D). To use the nomenclature common to proteases, the unattributed electron density extends as far as the S₁' subsite, the location ordinarily occupied by the scissile amino acid. The electron density in the S₁' subsite is of a size, shape, and location corresponding to that of alanine. The ligand is stabilized in this position by van der Waals' contacts between the alanyl methyl group

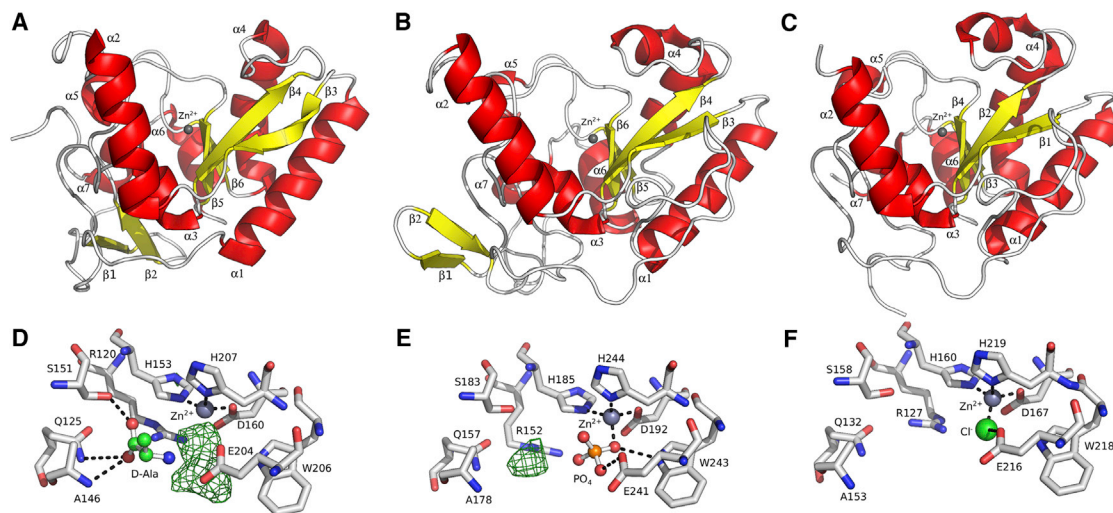


Figure 3. Structure of LdcB

(A–C) Cartoon representations of (A) *SpLdcB*, (B) *BsLdcB*, and (C) *BaLdcB*; α helices (red) and β strands (yellow) are numbered. The loops between secondary structure elements are colored silver. The bound zinc ion is shown as a gray sphere.

(D–F) The active sites of (D) *SpLdcB*, (E) *BsLdcB*, and (F) *BaLdcB*. A zinc ion is coordinated in a tetrahedral manner in the structures by two invariant histidines and a well-conserved aspartate. Residual electron density in *SpLdcB*, shown in (D), corresponds to D-alanine in the S_1' subsite, and an unknown compound in S_1 , adjacent to the zinc ion. Electron density corresponding to a phosphate ion, a component of the crystallization medium, in the S_1 subsite of *BsLdcB* shown in (E) is a mimic of the tetrahedral transition state. A second feature in the S_1' subsite is unmodeled. The residual difference density features are shown contoured at a level of 3σ and hydrogen bonds are drawn as dashed black lines.

and the phenolic side chains of Tyr191 and Tyr201 and by hydrogen bonds to the side chains of Glu204 and the main chain amide nitrogen of Tyr144. Despite the high resolution of the *SpLdcB* diffraction data, it was not possible to discriminate between either alanyl stereoisomer, both of which are present in the crystallization conditions. We have modeled D-Ala because it is the enzymatic product of these enzymes. However, residual electron density adjacent to the zinc ion could not be modeled satisfactorily with any ligand arising from the enzymatic activity, the crystallization media, or purification media; consequently, the remaining electron density has been left unmodeled. Mass differences between native and denatured protein samples also failed to identify candidate molecules corresponding to the residual electron density. Residual electron density in the vicinity of the fourth zinc ligand in *BsLdcB* was fulfilled by a phosphate anion, a component of the crystallization conditions (Figure 3E), which refined satisfactorily. The phosphate is located so as to mimic the tetrahedral transition state of amide bond hydrolysis. There is a further unmodeled electron density feature (Figure 3E), this time in the S_1' subsite, that we have not been able to identify unambiguously.

Comparison of LdcB Structures

SpLdcB, *BaLdcB*, and *BsLdcB* share pairwise sequence identities of ~30%, and they superimpose on ~155 common C_α atoms with a root-mean-square deviation (rmsd) of 1.5 Å–1.8 Å (Figures 3A–3C). The major differences between the LdcB structures are limited to the rather flexible and unconstrained N- and C-termini and the occasional longer loop in one structure in comparison to the other, as a direct consequence of sequence insertions. The active sites of *SpLdcB*,

BaLdcB, and *BsLdcB* are, to all extents and purposes, indistinguishable (Figures 3D–3F). A further *BaLdcB* structure has been solved that contains no bound zinc ion, but there are negligible differences between it and its counterpart with zinc bound (Figure S1 available online). The active site of the LdcB enzymes is, therefore, sculpted in a robust fashion, competent for peptidolysis once the catalytic machinery is completed by the incorporation of zinc.

Structural Comparison with Other Peptidases

Using *SpLdcB* as a search model in DALI (Holm and Rosenström, 2010), over 100 significantly similar structures were identified in the Protein Data Bank (PDB), including the structure of VanXY_G, a dipeptidase/DD-carboxypeptidase (PDB ID 4f78 [Meziane-Cherif et al., 2014]; DALI Z score = 17.0); an endolysin from the *Listeria* bacteriophage A500 (PDB ID 2vo9 [Korndörfer et al., 2008]; Z = 9.5); VanX, a DD-dipeptidase (PDB ID 1r44 [Bussiere et al., 1998]; Z = 6.6); and the N-terminal domain of sonic hedgehog, a protease involved in embryogenesis (PDB ID 2wg4 [Bishop et al., 2009]; Z = 5.3). Equivalent results were obtained using the structures of *BaLdcB* and *BsLdcB* as queries in DALI. The matched structures are all members of the wider LAS superfamily of proteases (Bochtler et al., 2004), which, although they share very little sequence identity, possess markedly similar active sites (Figure 4). By comparison with the VanXY_G structure (Meziane-Cherif et al., 2014), the active site of the LdcB proteins is open and more solvent accessible. The open nature of the active site is a direct consequence of the absence of the β_4 - α_5 /bisubstrate selectivity loop in LdcB, which enables the recognition of PG, a molecule far greater in size than the dipeptides utilized by VanXY_G (Meziane-Cherif et al., 2014).

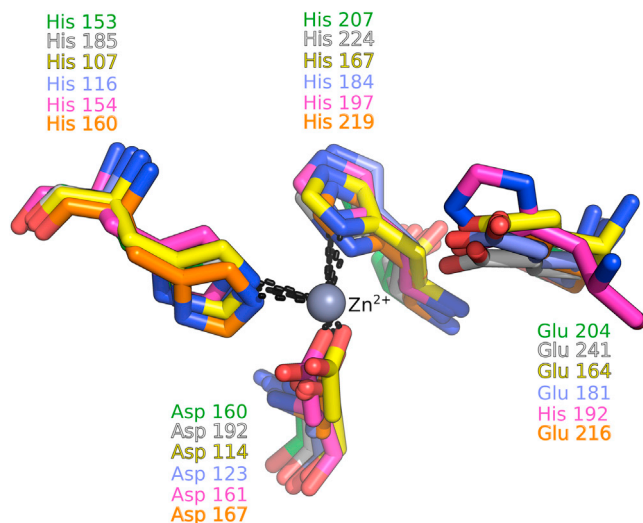


Figure 4. A Comparison of the Active Sites of LdcB and Other LAS Family Members

LdcB shares an active site closely superimposable on those of other LAS family enzymes. The zinc is coordinated by a $H-X_{(3-6)}-D-X_{(not\ conserved)}-H$ motif. A glutamate that plays a key role in catalysis is also shown; note that, in D-Ala-D-Ala carboxypeptidase from *Streptomyces albus*, the equivalent amino acid is histidine, the first amino acid in the $H-X-H$ motif common to most LAS carboxypeptidases. The structures shown, distinguished by the coloring of carbon atoms, are SpLdcB (white), BsLdcB (green), BaLdcB (orange), *Enterococcus faecalis* D,D-dipeptidase/D,D-carboxypeptidase VanXY_G (yellow; PDB ID 4f78), *Enterococcus faecium* aminodipeptidase VanX (blue; PDB ID 1r44) and *S. albus* D-Ala-D-Ala carboxypeptidase (pink; PDB ID 1lbu).

As with other LAS superfamily members, LdcB contains the characteristic zinc binding motif (Figure 4), with a consensus sequence of $H-X_{(3-6)}-D-X_{(not\ conserved)}-H$. The fourth zinc ligand in the structures of other LAS family members tends to be a solute from the crystallization milieu or a water that is likely to act as the nucleophile in the reaction (Bussiere et al., 1998). A second conserved motif in the LAS superfamily, $H-X-H$, of which the first histidine is the catalytic residue (Bochtler et al., 2004), is missing from LdcB enzymes; the catalytic histidine is replaced spatially by glutamate (i.e., Glu204 in SpLdcB).

SpLdcB Cocrystallized with a Reaction Product

Despite attempting to crystallise LdcB enzymes with a range of synthetic ligands, one crystal structure was obtained, that of SpLdcB in complex with MurNAc—L-Ala—D-γ-Gln—L-Lys—(D-Asn) (Figure 5A). This ligand is a close chemical analog of the true reaction product, as it contains half of the repeating unit of the glycan component of PG and comprises a tripeptide stem peptide that is branched at the L-lysine, the location of the crosslink position in natural PG (Figures 5F and 5G). There are two chemical differences between this synthetic mimic and the true reaction product. First, the ligand lacks GlcNAc, the second sugar moiety of the glycan repeating unit in PG. Second, the terminal D-Asn group is covalently attached by a peptide bond between the main chain carbonyl of the terminal D-Asn and the side chain N_{ϵ} of the penultimate lysine. In crosslinks between peptide stems, the side chain N_{ϵ} of the penultimate lysine is linked by a peptide bond to the main chain carbonyl of a

D-Ala. Hence, the L-Lys—(D-Asn) linkage is a faithful mimic of the natural L-Lys—(D-Ala), albeit the chemistry of the side chains of the terminal D-amino acids clearly differ.

There are six molecules of SpLdcB in this crystallographic asymmetric unit, though none of the interfaces represent a stable multimeric form. Five of six protein chains could be built and refined satisfactorily; the electron density for the sixth copy could not be improved by any procedure to permit its reliable modeling, and its absence from the model contributes directly to the slightly inflated final R_{work} and R_{free} values (Table 1). The five built molecules are highly superimposable, with each other and with the apo SpLdcB structure (rmsds between 0.5 Å and 1.1 Å; Table 2). In chains B, C, and E, the zinc is coordinated by a likely nucleophilic water to form tetrahedral coordination geometry. In chain D, the fourth zinc ligand is not evident, perhaps due to local disorder. Residual electron density was retained throughout refinement in only one of the five molecules, chain A, into which MurNAc—L-Ala—D-γ-Gln—L-Lys—(D-Asn) could be built and refined unambiguously (Figure 5A).

The structure of the ligand-bound form of SpLdcB is entirely consistent with its LD-carboxypeptidase function. First, the MurNAc moiety is situated at the very top of the intersubdomain cavity and makes little contact with the protein. This part of the ligand has the highest B factors (a mean of 93 Å², versus a mean of 83 Å² for the intact ligand and a mean of 45 Å² for all atoms modeled) and, thus, the weakest electron density. The MurNAc O1 and O4 atoms, which define the location of the flanking β1- to β4-linked GlcNAc sugars in PG, project away from the protein surface so that the glycan chain would be accommodated by the protein without steric clashes. The *N*-acetyl group at C1 is stabilized against the side chain of Met202 by van der Waals' forces (Figure 5B).

The C_{α} and C_{β} atoms of L-alanine at position 1 of the peptide pack against the face of Tyr144, a conserved amino acid in many LdcB sequences, to discriminate against the D-stereoisomer in this subsite, S_3 . L-alanine is favored by contacts with Tyr144 and Met202, which discriminate against all L-amino acids bulkier than alanine (Figure 5B). The NH₂ group of the D-γ-Gln of position 2 is located in subsite S_2 , 2.3 Å away from and in hydrogen bonding distance of Glu204 (Figure 5C), an invariant amino acid in LdcB sequences. Stereoselectivity at this position is maintained because L-γ-glutamine would clash with Tyr144. The γ-glutamine is selected for by specific interactions with the main chain carbonyl of Ala203 and the side chain of Glu204 while ensuring that the scissile peptide bond is in the catalytic position. These key aspects of molecular recognition could not occur with commonly occurring amino acids. Canonical glutamine can be accommodated at this position in the synthetic peptide, tetra-D, because the absence of contacts with one face of the peptide provides scope for the conformational changes needed to accommodate glutamine in the absence of the glycan strands of the natural substrate.

Position 3 of the peptide is occupied by L-lysine. D-lysine could not be accommodated at this position because of the presence nearby of Glu107. As expected of a reaction product complex, the lysine carboxylate is located adjacent to the active site zinc in subsite S_1 (Figure 5D). One of the carboxylic oxygens completes the zinc tetrahedral geometry with His153/Asp160/His207, and the second oxygen forms a salt bridge with

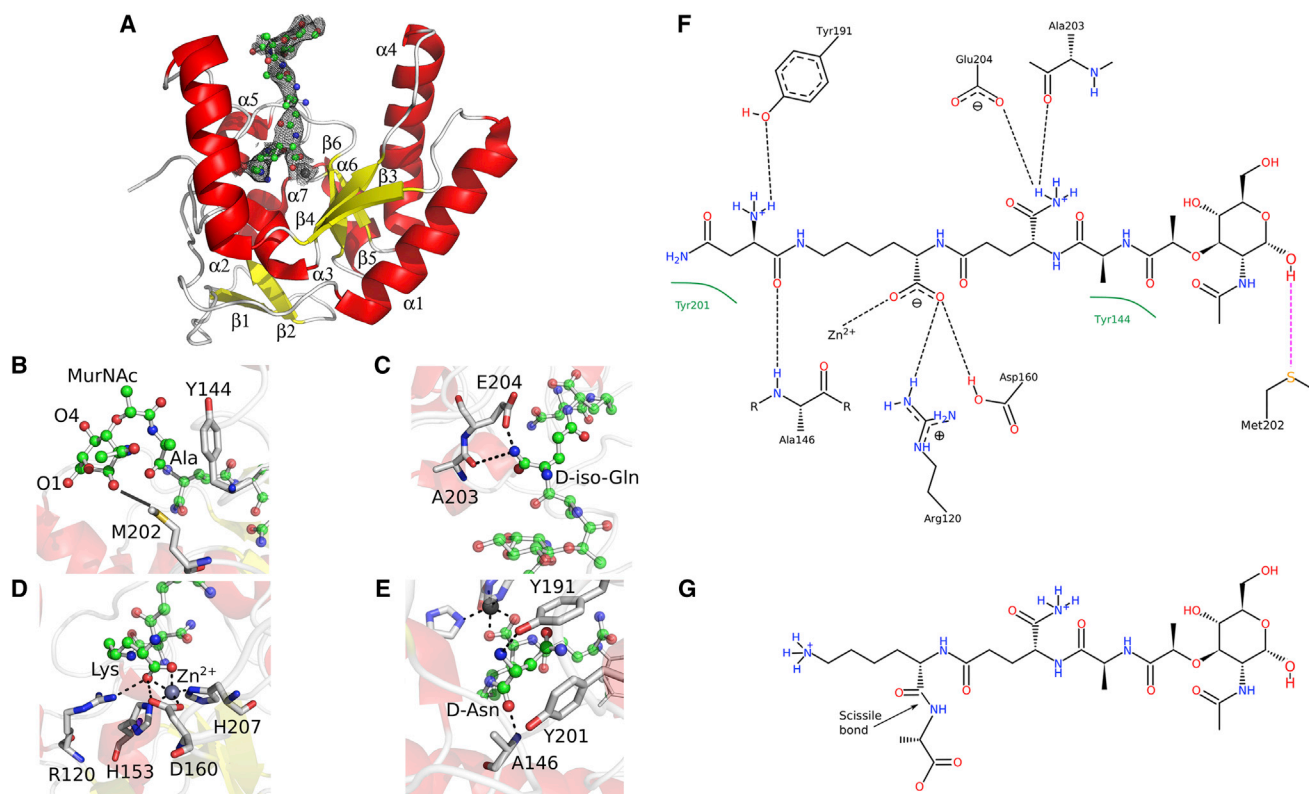


Figure 5. SpLdcB with Bound MurNAc—L-Ala—D-γ-Gln—L-Lys—(D-Asn)

(A) Cartoon representation of SpLdcB, drawn using the same color scheme as in Figure 3. The MurNAc—Ala—D-γ-Gln—L-Lys—(D-Asn) is drawn as a ball-and-stick model with Refmac-weighted $2F_{\text{obs}} - F_{\text{calc}}$ electron density displayed at a contour level of 1σ .

(B–E) The interactions made between the ligand and the protein at each subsite are shown successive panels: (B) S_3 ; (C) S_2 ; (D) S_1 , and (E) S_1' . Key van der Waals' interactions are shown as a transparent silver line.

(F) A schematic of the interactions of SpLdcB with the bound peptidoglycan mimic. Hydrogen-bond interactions are shown with a black dotted line, and van der Waals' forces are shown with a green arc.

(G) A schematic of the tetrapeptide substrate of SpLdcB, drawn in the same manner as the peptidoglycan mimic in (F). The scissile bond is highlighted by an arrow. Crosslinks to other stem peptides would take place via the lysine's terminal amino group on the left.

Arg120, which is either arginine or lysine in LdcB orthologs. The L-lysine carboxylate thus represents a posthydrolysis state, placed appropriately for the peptide bond of the substrate to be hydrolyzed by the water molecule that fulfils the fourth zinc ligand in the absence of substrate.

The side chain of the L-lysine twists back against the rest of the peptide and projects into the protein core to place the terminal D-Asn branch in subsite S_1' , against the main chain atoms of Tyr144 and the side chains of Tyr191, Tyr201, and His207 (Figure 5E). The carbonyl of the L-Lys—(D-Asn) peptide bond is contacted by the side chain of Ser151. However, it should be noted that subsite S_1' would be occupied by the scissile D-alanine in the presence of the reaction substrate, and support for this hypothesis comes from the structure of VanXY_G solved in complex with a phosphinate transition state mimic (Figure 6A) (PDB ID 4muq) (Meziane-Cherif et al., 2014). The phosphinate moiety is a transition state mimic of peptide bond hydrolysis because the reaction proceeds via a tetrahedral intermediate that phosphinate mimics. VanXY_G superimposes on SpLdcB with an rmsd of 1.6 Å on 144 matched C $_{\alpha}$ s, and the alanyl carboxylate moiety of the phosphinate is situated in an analogous position

to that of the L-Lys—(D-Asn) peptide linkage observed in SpLdcB (Figure 6A). The phosphate atoms of the phosphinate transition state mimic phosphinate almost identically with the phosphate anion modeled adjacent to the zinc ion in the structure of BsLdcB (Figure 3D). Other than Ser151, a role in determining specificity for the terminal D-alanine can be invoked for Gln125, and both these amino acids are strictly conserved. The amide nitrogen of Ala146 also participates in stereoselectivity, but as this is a main chain atom, it is perhaps no surprise that sequence conservation is not maintained at this position.

However, crosslinks between peptide stems cannot be accommodated in the observed conformation because of the burial of the branched amino acid in subsite S_1' (Figure 6B). Mere rotation of the side chain torsion angles of the L-lysine would result in a more thermodynamically favored extended conformation for this side chain to result in it passing beyond Arg120, Leu128, and Tyr132 to project beyond the surface of the protein (Figure 6B; Figure S2). In this conformation, crosslinked peptide stems attached to the terminal N $_{\epsilon}$ would not be sterically occluded by the body of the protein, consistent with the observed cleavage of crosslinked peptides (Figure 2A). In

Table 1. Summary of X-Ray Data Collection and Refinement Statistics

Parameters	SpLdcB	BsLdcB	SpLdcB with Ligand	BaLdcB	BaLdcB with Zinc
PDB ID	4OX5	4OX3	4OXD	4JID	4MPH
Data collection					
Space group	<i>I</i> 222	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>C</i> 2	<i>I</i> 222	<i>I</i> 222
Unit Cell Dimensions					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	48.8, 60.4, 138.2	42.7, 53.7, 102.4	346.0, 42.5, 79.3	76.6, 113.0, 124.6	76.8, 112.3, 125.7
α , β , γ (°)	90, 90, 90	90, 90, 90	90, 93.1, 90	90, 90, 90	90, 90, 90
Wavelength (Å)	0.98	0.97	0.92	0.97856	1.27696
Resolution (Å)	20.0–1.80 (1.84–1.80)	47.6–1.70 (1.73–1.70)	47.8–2.80 (2.95–2.80)	29.7–2.30 (2.36–2.30)	28.3–2.03 (2.06–2.03)
Multiplicity	24.3 (14.2)	3.2 (1.7)	3.4 (3.4)	4.2 (3.6)	11.8 (9.9)
<i>R</i> _{merge} ^a	0.077 (0.108)	0.072 (0.235)	0.134 (0.635)	0.093 (0.542)	0.101 (0.511)
<i>I</i> / σ <i>I</i>	34.4 (2.4)	11.9 (2.4)	7.8 (1.9)	13.5 (2.4)	32.1 (4.2)
Completeness (%)	98.6 (86.9)	95.7 (89.3)	98.3 (98.3)	98.3 (94.3)	99.8 (96.9)
Refinement					
Resolution (Å)	20.0–1.80	47.6–2.0	47.8–2.80	29.7–2.3	28.3–2.03
No. of reflections	25,656 (1,157)	35,526 (1,057)	47,327 (1,979)	22,730 (1,226)	67,494 (3,374)
<i>R</i> _{work} / <i>R</i> _{free} ^b	0.191/0.253	0.176/0.247	0.273/0.335	0.194/0.259	0.151/0.189
No. of Atoms					
Protein	1,476	1,753	7,014	3,040	2,970
No. of proteins/AU	1	1	6	2	2
Ion/ligand	3/22	1/5	21/51	–/4	2/12
Water	197	230	111	424	466
B Factors (Å ²)					
Protein	41.8	16.0	44.3	45.3	35.8
Ion/ligand	36.9/50.7	10.7/23.8	38.7/69.7	–/49.8	17.9/60.0
Water	48.2	31.6	28.7	43.6	46.2
Rmsd					
Bond lengths (Å)	0.018	0.015	0.005	0.013	0.0012
Bond angles (°)	1.96	1.76	0.84	1.36	0.991
Ramachandran					
Favored (%)	96.1	96.7	95.8	97.6	98.0
Allowed (%)	100	100	99.6	99.7	100

Highest resolution shell shown in parentheses. AU, asymmetric units.

^a*R*_{merge} = $\sum_i \sum_h |I_i(h) - \langle I(h) \rangle| / \sum_i \sum_h I_i(h)$, where *I*_{*i*}(*h*) and $\langle I(h) \rangle$ are the *i*th and mean measurement of the intensity of reflection *h*.

^b*R*_{work}/*R*_{free} = $\sum |F_p^{obs} - F_p^{calc}| / \sum F_p^{obs}$, where *F*_{*p*}^{obs} and *F*_{*p*}^{calc} are the observed and calculated structure factor amplitudes, respectively.

some bacteria, the lysine is replaced with diaminopimelic acid, a carboxy derivative of lysine that is amidated in bacilli. While the amidocarboxy derivative could be stabilized in the observed conformation by a similar set of specific interactions, it would also sterically occlude crosslinked peptide stems. Diaminopimelic acid is thus likely to follow the same path predicted for lysine by forming a simple, extended conformation to project into solvent.

Several key structural changes occur on ligand binding in SpLdcB (Figure 7; Movie S1). First and foremost, substrates are occluded in the empty form of the enzyme by a loop between residues Gly163 and Glu171, which blocks access to the active site. Most of these residues (Asp167 to Glu171) move by 10–15 Å in the presence of ligand to open up the cleft by refolding into an additional turn at the N terminus of α -helix 4. The rearrangement of these residues is accompanied by further conformational changes in Trp206, the indole ring of which moves by up to 9 Å,

which not only aids to open up the active site of the enzyme but also forces the side chain of Glu204 to rearrange to take up a position capable of interacting with the substrate. The importance of Glu204 and the zinc ion for catalysis is highlighted by the observation that tetra-D is a much poorer substrate for the E204A variant of SpLdcB (Figure 2A) and by the absence of activity in the presence of EDTA (Figures 2A and 2C). The conformational changes seen on ligand binding to SpLdcB contrasts with the absence of significant conformational changes described for VanXY_G in complex with phosphinate (PDB ID 4muq) (Meziane-Cherif et al., 2014). Like LdcB, VanXY_G maintains the H-X_(3–6)-D-X_(not conserved)-H motif, and as also observed in LdcB, the catalytic H-X-H motif is replaced by a glutamate. The resculpting of the N terminus of α -helix 4 in SpLdcB to accommodate the tetrapeptide is not necessary in VanXY_G, because this enzyme hydrolyzes a much smaller substrate, D-Ala-D-Ala dipeptide, which is accommodated by the enzyme without significant rearrangement.

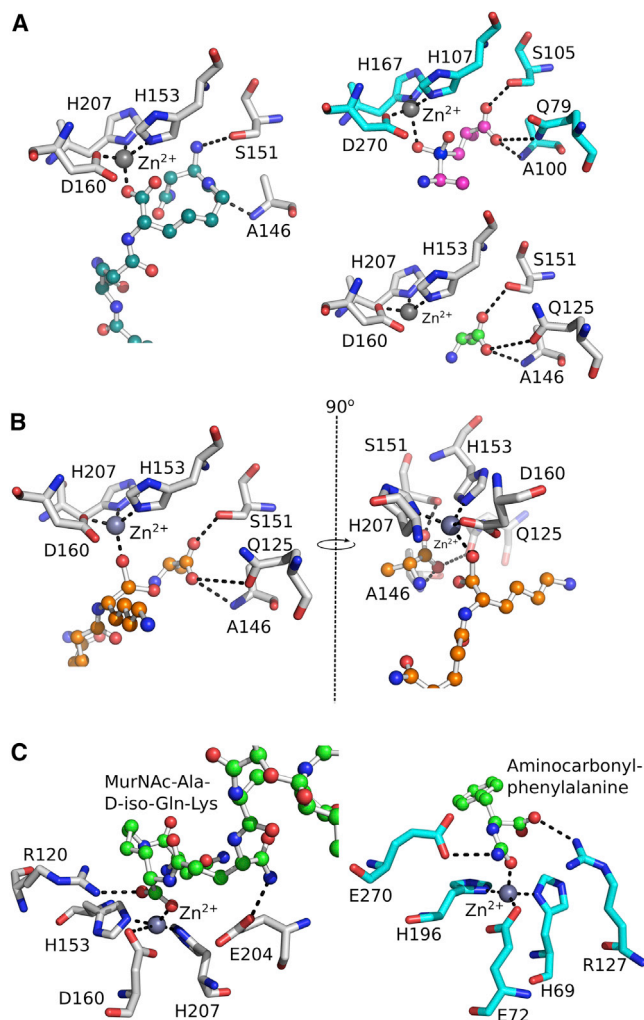
Table 2. Rmsd Values between Ligand-Bound SpLdcB Chains (4OXD) and apo SpLdcB (4OX5)

4OX5 and 4OXD Chain	4OX5	4OXD Chain A	4OXD Chain B	4OXD Chain C	4OXD Chain D	4OXD Chain E
4OX5	–	0.73	0.57	0.66	0.69	0.80
4OXD chain A	0.73	–	0.63	0.64	0.76	1.03
4OXD chain B	0.57	0.63	–	0.56	0.65	0.86
4OXD chain C	0.66	0.64	0.56	–	0.56	0.64
4OXD chain D	0.69	0.76	0.65	0.56	–	0.61
4OXD chain E	0.80	1.03	0.86	0.64	0.61	–

DISCUSSION

The presence of tripeptides in the bacterial cell wall implies that LD-carboxypeptidase activity must be present in the extracellular milieu. In this study, we have shown that LdcB, from both *S. pneumoniae* and *B. subtilis*, are active against tetrapeptides found within the bacterial cell wall and are specific for LD linkages, confirming the LD-carboxypeptidase classification with a role in cell wall maturation. The precise *in vivo* role for LdcB remains to be determined; one possibility is that the enzyme allows recycling of D-Ala from the cell wall to support further cell wall synthesis and, thus, growth. Alternatively, the cleavage of D-Ala from the cell wall might permit the differentiation between old and new cell walls, a potentially powerful way to ensure the correct localization of new cell wall synthesis. The correct localization of new PG synthesis is likely to be more important to bacteria that have polarized cell growth, such as *Streptococci*, in comparison to those that have dispersed growth along the length of a cylinder (e.g., *Bacilli*) or that utilize a division-only mechanism, such as *Staphylococci*. The cell morphological defect observed in the *S. pneumoniae* LdcB mutant, and the absence of defect in the *B. subtilis* LdcB mutant, could be consistent with this premise. Further studies should distinguish these possibilities.

We also present crystal structures of BaLdcB, BsLdcB, and SpLdcB, the latter also in complex with a PG-fragment analog that has enabled the identification of amino acids key to molecular recognition of the substrate and a potential tunnel through which crosslinked PG peptides may pass during proteolysis. LdcB belongs to the LAS group of zinc metalloproteases; several members of this family are active against PG, albeit with different substrate specificities to LdcB, including the Gly-Gly endopeptidase lysostaphin (Korndörfer et al., 2008) and the DD-/LD-endopeptidase MepA (Marcyjaniak et al., 2004). LdcB is dissimilar to other groups of LD-carboxypeptidases, e.g., *Pseudomonas aeruginosa* LD-carboxypeptidase LdcA (PDB ID 1ZRS) (Korza and Bochtler, 2005) that degrades cell wall tetrapeptides in the cytoplasm. LdcA utilizes a Ser-His-Glu catalytic triad in catalysis and belongs to the S66 serine protease family of enzymes, a group with which LdcB shares no similarity (Korza and Bochtler, 2005). Likewise, LdcB shares no sequence similarity with Csd6 and Pgp2, two recently discovered LD-carboxypeptidases that aid to maintain the helical shape of *H. pylori* (Sycuro et al., 2012) and *C. jejuni* (Friedrich et al., 2014). Csd6 and Pgp2 belong to an LD-transpeptidase family that utilizes a conserved Cys/His motif for catalysis (Biarrotte-Sorin et al., 2006), and are,

**Figure 6. A Comparison of the Ligand-Bound Active Sites Carboxypeptidases**

(A) The active site of SpLdcB (white) with bound MurNAc-L-Ala-D-γ-Gln-L-Lys-(D-Asn) (turquoise carbons) is compared to that of VanXYg (blue) containing a phosphinate transition state analog (pink carbons) of D-Ala-D-Ala (PDB ID 4muq). The phosphinate in VanXYg coordinates the active site zinc in a similar manner to that of the lysyl carboxylate in the MurNAc-L-Ala-D-γ-Gln-L-Lys-(D-Asn) bound to SpLdcB. The D-alanyl moiety of the phosphinate ligand in VanXYg, which occupies subsite S₁', is matched by the D-alanine (green carbons) in SpLdcB (white) and the D-Asn of the MurNAc-L-Ala-D-γ-Gln-L-Lys-(D-Asn) ligand. The zinc ions are shown as gray spheres.

(B) The proposed positions of the two posthydrolysis ligands (orange carbons) in the active site of SpLdcB (white). The D-alanine from the SpLdcB structure occupies the S₁' site of the protein, while the lysine side chain has been rotated and modeled in an extended conformation. The altered conformation of the lysine presents the N_ε to the surface of the protein, in an ideal position to accommodate peptide crosslinks. The zinc ions are shown as gray spheres.

(C) The active sites of SpLdcB and pancreatic carboxypeptidase A (PDB ID 1hdu [Cho et al., 2002]) are compared. In each structure, the Zn²⁺ ion is coordinated by two histidines (His69 and His196 in carboxypeptidase A; His153 and His207 in SpLdcB) and an acidic amino acid (Glu72 in carboxypeptidase A; Asp160 in SpLdcB). The reversed locations of Arg127 and Glu270 in carboxypeptidase A relative to Arg120 and Glu204 in SpLdcB is matched by a reversal of the path of the respective substrates through the active sites and the likely retention of a common catalytic mechanism.

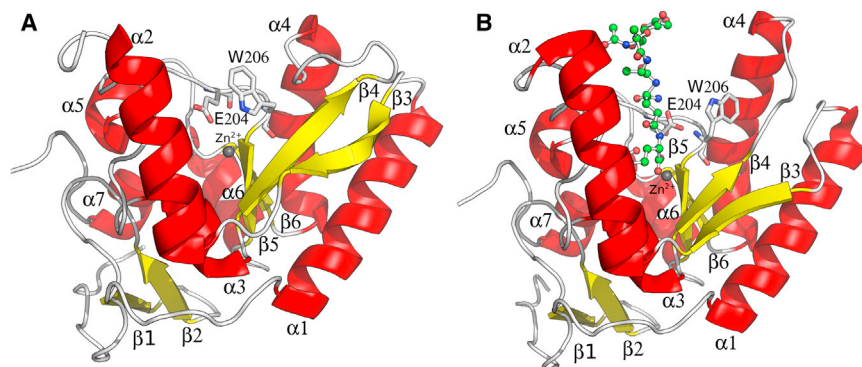


Figure 7. Conformational Changes in SpLdcB on Ligand Binding

(A and B) Cartoon representations of (A) apo SpLdcB and (B) SpLdcB with bound MurNAc-Ala-D-γ-Gln-Lys-(D-Asn) using the same color scheme as in Figure 3. Key residues that move on ligand binding are shown in stick representation.

therefore, distinct from the metalloprotease LdcB. It is interesting that Csd6 can only hydrolyze D-Ala from monomeric stem peptides, whereas LdcB is active against both monomeric and crosslinked (tetra-tetra) stem peptides (Sycuro et al., 2013), but the absence of structural similarity between these enzymes precludes useful comparison.

While LdcB shares no structural homology with previously studied LD-carboxypeptidases, DALI searches revealed a strong similarity to members of the LAS superfamily. Of the two consensus sequences that characterize the LAS family, LdcB lacks the second His-rich motif, which is replaced spatially by the invariant Glu204, which is oriented away from the zinc ion. Given the residual activity of the E204A mutant (Figure 2A), Glu204 probably plays an indirect role in catalysis. The zinc ion probably plays a more direct role in catalysis, a hypothesis that is supported by the absence of enzymatic activity in the presence of EDTA (Figures 2A and 2C). Further support for the role of the metal ion, and the indirect role for the glutamate, comes from studies of other zinc-dependent metalloproteases such as pancreatic carboxypeptidase A, which is of historical importance given its role in the demonstration of Koshland's induced fit hypothesis (Christianson and Lipscomb, 1989; Koshland, 1958). Given the conformational changes seen on binding ligand (Figure 7; Movie S1), LdcB appears to be another example of induced fit. Though the two enzymes are structurally distinct, the active sites of LdcB and pancreatic carboxypeptidase A are, in essence, mirror images of one another. Both contain a zinc ion coordinated by two histidines (His69 and His196 in carboxypeptidase A; His153 and His207 in SpLdcB) and an acidic amino acid (Glu72 in carboxypeptidase A; Asp160 in SpLdcB). SpLdcB Arg120 and Glu204 straddle the nucleophilic water, but their positions relative to their counterparts in pancreatic carboxypeptidase A (Arg127 and Glu270) are reversed (Figure 6C). Consequently, the path of the substrate through the active site is also reversed. Given the retention of the metal coordination and key adjacent amino acids in LdcB, it is entirely plausible that the reaction mechanism of LdcB occurs in a similar fashion to that of pancreatic carboxypeptidase A, utilizing the metal ion to activate the nucleophilic water to attack the scissile peptide bond. The enzymes would thus appear to have converged on a common mechanism, despite sharing little overall sequence and structural homology. Confirmation of the catalytic mechanism utilized by this enzyme, and its precise role in vivo, is the focus of current research activities in our laboratories.

EXPERIMENTAL PROCEDURES

Cloning

B. subtilis LdcB and *S. pneumoniae* LdcB were amplified by PCR from *B. subtilis* 168 and *S. pneumoniae* R6 genomic DNA, respectively, to generate fragments with NdeI and XhoI restriction sites. Both fragments were digested with NdeI and XhoI and ligated with NdeI/XhoI-digested pET28a. *B. anthracis* LdcB was PCR amplified from *B. anthracis* str. Ames ancestor genomic DNA and cloned into p15TV-LIC (Eschenfeldt et al., 2009) that codes for a N-terminal His₆ tag, followed by a tobacco etch virus (TEV) protease cleavage site. In all cases, the recombinant constructs were designed to remove membrane-anchoring amino acids to result in final clones encoding SpLdcB amino acids 56–238; BsLdcB, 61–273; and BaLdcB, 55–243.

Strain Construction

S. pneumoniae R6ΔLdcB was constructed by transforming strain R6 with a PCR product covering the LdcB-deletion region from strain IU3880 (D39 Δcps2A'(cps2BCDETFG)H'ΔdacB <> P_c-kan-rpsL) amplified with primers P136 and P137 (Barendt et al., 2011), followed by selection on tryptic soy broth blood agar plates containing 250 μg/ml kanamycin. All strains were grown in C + Y medium (Horne and Tomasz, 1993), containing kanamycin if required. Competent bacteria were obtained as described elsewhere, with the addition of competence-stimulating peptide (Vollmer and Tomasz, 2001).

B. subtilis strains were constructed by transformation of chromosomal DNA (Anagnostopoulos and Spizizen, 1961; Young and Spizizen, 1961) from strains bearing null mutations for *dacA* (MC01, *trpC2* Δ*dacA cat*) or *yodJ* (MC05, *trpC2* Δ*yodJ neo*) into strain RD180 (*trpC2* Δ*alrA zeo*), selecting for the appropriate resistance markers (chloramphenicol or kanamycin at 5 μg/ml) in the presence of 40 μg/ml D-alanine, yielding strains KS03 and KS02, respectively.

Growth Curves

B. subtilis strains RD180, KS02, and KS03 were grown overnight at 30°C in Luria broth (LB) supplemented with 40 μg/ml D-Ala. The overnight cultures were diluted 1:10 with LB supplemented with 40 μg/ml D-Ala and grown at 37°C for 2 hr. The cells were collected by centrifugation and resuspended in D-Ala-free media at 37°C. This suspension was used to inoculate media with various concentrations of D-Ala to an optical density 600 (OD₆₀₀) of 0.05, and the subsequent growth of each culture at 37°C was monitored by determining the optical density at 600 nm every 20 min for 3 hr. The growth rate was determined during the exponential growth phase only.

Mutagenesis

A modified QuikChange protocol, incorporating “touchdown” (Don et al., 1991), was used for site-directed mutagenesis. The annealing temperature of the first 10 cycles was reduced by one degree per cycle, from 68°C to 58°C, before a further ten cycles were performed at a constant annealing temperature of 58°C. The extension temperature was maintained at 72°C for 11 min per cycle. Mutagenic primers (0.3 μl) were mixed with 50 ng of template DNA, 0.5 mM dinucleotide triphosphates, and 2.5 U of *Pfu* DNA polymerase in a total reaction volume of 20 μl. The template plasmid was digested using DpnI.

Peptidoglycan Activity Assays

Cell wall and PG from *S. pneumoniae* and *B. subtilis* strains were isolated by published procedures (Bui et al., 2012; Garcia-Bustos et al., 1988). To test for LD-carboxypeptidase activity, PG from *S. pneumoniae* R6ΔLdcB

(1.6 mg/ml) or *B. subtilis* 168Δ*ldcB* (3.1 mg/ml) was incubated in a total volume of 160 μl of 20 mM sodium phosphate, pH 4.8, with 0.13 mg/ml SpLdcB and 0.31 mg/ml BsLdcB, respectively, for 16 hr at 37°C. If required, 0.6 mM Zn²⁺ or 10 mM EDTA was added, and control samples contained no added enzyme. The PG was digested with 25 μg/ml of the muramidase cellosyl (a kind gift of Hoechst AG) for 24 hr at 37°C. The samples were boiled for 10 min and centrifuged for 10 min at 15,000 × g. The muropeptides present in the supernatant were reduced with sodium borohydride and analyzed by high-pressure liquid chromatography (HPLC) (Bui et al., 2012; Garcia-Bustos et al., 1988). *S. pneumoniae* and *B. subtilis* muropeptide peaks were assigned based on their published elution profile and retention times (Bui et al., 2012; Atrih et al., 1999). Fractions containing the reduced muropeptides Tetra and TetraTetra present in *B. subtilis* strain 168Δ*ldcB* were collected and analyzed by mass spectrometry (Bui et al., 2009). We determined a neutral mass of 940.4353 atomic mass units (amu) for Tetra (theoretical: 940.4237 amu) and of 1862.8406 amu for TetraTetra (theoretical: 1862.8368 amu).

Thin-Layer Chromatography

Tetrapeptides Ala-D-Gln-L-Lys-D-Ala (Tetra-D; Activotec) or Ala-D-Gln-L-Lys-L-Ala (Tetra-L; Activotec) at 20 mM were incubated with 700 μM of zinc-loaded SpLdcB or SpLdcB-E153A for 2 hr at 37°C in a buffer of 10 mM HEPES/NaOH, pH 8.0, 100 mM NaCl in a reaction volume of 10 μl. One sample of SpLdcB contained 50 mM EDTA. Ethanol (5 μl) was added to the reactions before spotting on to silica gel thin-layer chromatography (TLC) aluminum foils (Sigma). The TLCs were developed in a running buffer of butan-1-ol, acetic acid, water in a 3:1:1 ratio for 40 min before staining with 0.1% ninhydrin solution in ethanol (Sigma). The TLC foils were dried and warmed with a domestic hairdryer for visualizing the results.

Protein Overexpression and Purification

BsLdcB and SpLdcB were prepared in the same way. The proteins were overexpressed in the methionine auxotrophic B834(DE3) strain of *E. coli* by growth in LB media (or in minimal medium supplemented with selenomethionine for selenomethionine-labeled protein) at 37°C up to an OD₆₀₀ of 0.4 before induction with 1 mM IPTG for 18 hr at 16°C. Cells were harvested by centrifugation at 3000 g for 25 min at 4°C and the pellet was resuspended in 40 ml buffer A (50 mM HEPES/NaOH, pH 8.0, 300 mM NaCl, 20 mM imidazole) per liter of culture and lysed on ice for 3 min with a Sonopuls HD2070 (Bandelin) sonicator. The cell lysate was clarified at 33000 g for 20 min and the supernatant loaded onto a 5 ml Ni²⁺-NTA column (GE Healthcare) pre-equilibrated with buffer A. The column was washed with buffer A before elution of bound proteins with buffer B (50 mM HEPES/NaOH, pH 8.0, 300 mM NaCl, 500 mM imidazole). The fractions containing LdcB were pooled and the N-terminal hexahistidine tag was removed by digestion with thrombin (Sigma-Aldrich) at 20°C for 16 hr at a ratio of 1 unit of enzyme per mg of LdcB. LdcB was concentrated to a volume of 1 ml before loading onto a Superdex S75 HiLoad 16/60 (GE Healthcare) gel filtration column pre-equilibrated in buffer C (10 mM HEPES/NaOH, pH 8.0, 100 mM NaCl).

For BaLdcB, freshly transformed *E. coli* BL21-CodonPlus (DE3)-RIPL cells were grown in 0.5 l of LB medium containing 100 μg/ml ampicillin. Cultures were grown at 37°C to an OD₆₀₀ of 0.8 and induced for 16 hr at 16°C with 1 mM IPTG, and the cells were harvested by centrifugation and resuspended in 20 ml of buffer A (50 mM HEPES, pH 7.5, 300 mM NaCl, 10% [v/v] glycerol) supplemented with 10% (v/v) BugBuster 10× Protein Extraction Reagent (Novagen), 25 U Benzamide (Sigma-Aldrich), and 5 mM imidazole. The mixture was stirred for 10 min at room temperature and centrifuged at 20,000 × g for 45 min, and the supernatant was applied to a 1 ml HisTrap Fast Flow column (GE Healthcare) equilibrated with buffer A containing 40 mM imidazole. The protein was eluted with buffer A with a gradient of 40–500 mM imidazole over 20 ml. Fractions containing the recombinant His-tagged proteins were identified by SDS-PAGE; pooled; dialyzed overnight against 3 l of 50 mM HEPES, pH 7.5, 300 mM NaCl, 5% (v/v) glycerol, 1 mM TCEP; and stored at −80°C. The His₆-tag was cleaved by TEV protease in 500 mM NaCl, 50 mM Tris-HCl, pH 7.5, 1 mM TCEP, 1 mM EDTA, 20% glycerol.

Crystallization

BsLdcB and SpLdcB were concentrated to 15 mg/ml, and two molar equivalents of zinc chloride were added. Crystallization conditions for BsLdcB

and SpLdcB were obtained by sitting drop vapor diffusion at 20°C. The best SpLdcB crystals were obtained from a 0.1 M mixture of amino acids (glycine, L-glutamate, and racemates of alanine, serine, and lysine), 0.1 M Morpheus buffer system 2, pH 7.5, and 37.5% 2-Methyl-2,4-pentanediol (MPD)/polyethylene glycol (PEG) 1000/PEG 3350 (Gorrec, 2009). A single BsLdcB crystal grew from 0.1 M phosphate/citrate buffer, pH 4.2, 40% PEG 300 over a period of ~2 months. Crystals of SpLdcB with bound MurNAc-L-Ala-D-γ-Gln-L-Lys-(D-Asn) (InvivoGen) were obtained by incubating the protein with the ligand at 2 mM final concentration for 30 min at room temperature prior to crystallization against a mother liquor of 0.1 M ammonium chloride, 0.1 M HEPES/NaOH, pH 8.0, and 20% PEG 6000. Crystals of BaLdcB (zinc free) were grown at 23°C using hanging-drop vapor diffusion by mixing 20 mg/ml protein with reservoir solution containing 0.3 M sodium chloride, 0.1 M HEPES, pH 7.5. As soaking of zinc chloride into BaLdcB apo-protein crystals did not provide sufficient occupancy of zinc into the active site, crystals of BaLdcB (zinc bound) were grown at 23°C by hanging-drop vapor diffusion by mixing 20 mg/ml protein and 5 mM zinc chloride prior to crystallization with 0.2 M ammonium acetate, 0.1 M sodium acetate, and 30% PEG 4000.

Data Collection and Processing

All crystals were cryoprotected with paratone-N before flash cooling in liquid nitrogen prior to data collection at 100 K. Diffraction data for BsLdcB and SpLdcB were on beamline IO2 of the Diamond Light Source. For selenomethionine-labeled SpLdcB, diffraction data were collected at a wavelength of 0.98 Å to permit phasing by selenomethionine-SAD (Se-SAD). Diffraction data were integrated with XDS (Kabsch, 2010) and reduced with SCALA (Evans, 2006). For BaLdcB, diffraction data were collected at on beamlines 21-ID-G at Life Sciences Collaborative Access Team, Advanced Photon Source, at 0.97856 Å (selenomethionine peak; zinc free) and 1.27696 Å (Zn²⁺ peak; zinc bound). The BaLdcB diffraction data were integrated and reduced with HKL-3000 (Minor et al., 2006).

Structure Solution

The structure of SpLdcB was solved using SHELX (Sheldrick, 2008) as implemented in the HKL2MAP interface (Pape and Schneider, 2004). An initial model was built by PHENIX.AUTOSOLVE (Adams et al., 2011). The completed SpLdcB structure was used as the molecular replacement search model in PHASER (McCoy et al., 2007) to solve BsLdcB and the SpLdcB-ligand complex. The BaLdcB Zn²⁺-bound structure was solved by SAD phasing using PHENIX.solve (Adams et al., 2011), which identified the positions of the two bound Zn²⁺ ions. The occupancies of the zinc atoms were refined during model refinement. The BaLdcB Zn²⁺-free structure was solved by molecular replacement using the D,D-peptidase domain of the apo structure of VanXY_G (PDB ID 4f78) (Meziane-Cherif et al., 2014).

For all structures, manual building cycles in COOT (Emsley et al., 2010) were interspersed with restrained refinement in PHENIX.REFINE (Afonine et al., 2012). B factors were refined isotropically for all structures apart from BsLdcB, which was refined anisotropically. Noncrystallographic symmetry refinement was not used. Average B factor and geometric parameters were calculated using PHENIX and verified with PHENIX.refine and the Research Collaboratory for Structural Bioinformatics PDB Validation server. Data collection and final model refinement statistics for all structures can be found in Table 1. The graphics programs UCSF Chimera and PyMOL were used to generate all molecular figures presented.

ACCESSION NUMBERS

The PDB accession numbers for SpLdcB, BsLdcB, SpLdcB with bound ligand, apo BaLdcB, and BaLdcB with bound zinc reported in this paper are 4OX5, 4OX3, 4OXD, 4JID, and 4MPH, respectively.

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and one movie and can be found with this article online at <http://dx.doi.org/10.1016/j.str.2014.04.015>.

ACKNOWLEDGMENTS

We thank the staff at the Diamond Synchrotron Light Source for access to, and help with, its beamlines for X-ray diffraction data collection. We thank Arnaud Baslé, Joseph Newman, and Vincent Rao for useful discussions; Atikah Mohd Sukor for excellent technical assistance; Z. Wawrzak, S. Shatsman, and S.N. Peterson for assistance with data collection and refinement; and T. Skarina for assistance with purification and crystallization. We also thank the authors of [Meziane-Cherif et al. \(2014\)](#) for sharing results and atomic coordinates prior to publication. Finally, we thank Joe Gray of the Newcastle University Pinnacle-Proteomics and Biological Mass Spectrometry facility for mass spectrometry analysis of mucopeptide fractions. This project has been funded in whole or in part with federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, U.S. Department of Health and Human Services (contracts HHSN272200700058C and HHSN2722012000026C), and by the UK Biotechnology and Biological Sciences Research Council in an award to R.A.D., W.V., and R.J.L. (BB/G015902/1).

Received: March 4, 2014

Revised: April 24, 2014

Accepted: April 29, 2014

Published: June 5, 2014

REFERENCES

- Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Echols, N., Headd, J.J., Hung, L.W., Jain, S., Kapral, G.J., Grosse Kunstleve, R.W., et al. (2011). The Phenix software for automated determination of macromolecular structures. *Methods* 55, 94–106.
- Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* 68, 352–367.
- Anagnostopoulos, C., and Spizizen, J. (1961). Requirements for transformation in *Bacillus subtilis*. *J. Bacteriol.* 81, 741–746.
- Arminjon, F., Guinand, M., Vacheron, M.J., and Michel, G. (1977). Specificity profiles of the membrane-bound γ -D-glutamyl-(L)-meso-diaminopimelate endopeptidase and LD-carboxypeptidase from *Bacillus sphaericus* 9602. *Eur. J. Biochem.* 73, 557–565.
- Atrih, A., Bacher, G., Allmaier, G., Williamson, M.P., and Foster, S.J. (1999). Analysis of peptidoglycan structure from vegetative cells of *Bacillus subtilis* 168 and role of PBP 5 in peptidoglycan maturation. *J. Bacteriol.* 181, 3956–3966.
- Barendt, S.M., Sham, L.T., and Winkler, M.E. (2011). Characterization of mutants deficient in the L,D-carboxypeptidase (DacB) and WalRK (VicRK) regulon, involved in peptidoglycan maturation of *Streptococcus pneumoniae* serotype 2 strain D39. *J. Bacteriol.* 193, 2290–2300.
- Biarrotte-Sorin, S., Hugonnet, J.E., Delfosse, V., Mainardi, J.L., Gutmann, L., Arthur, M., and Mayer, C. (2006). Crystal structure of a novel β -lactam-insensitive peptidoglycan transpeptidase. *J. Mol. Biol.* 359, 533–538.
- Bishop, B., Aricescu, A.R., Harlos, K., O'Callaghan, C.A., Jones, E.Y., and Siebold, C. (2009). Structural insights into hedgehog ligand sequestration by the human hedgehog-interacting protein HHIP. *Nat. Struct. Mol. Biol.* 16, 698–703.
- Bochtler, M., Odintsov, S.G., Marcyjaniak, M., and Sabala, I. (2004). Similar active sites in lysostaphins and D-Ala-D-Ala metallopeptidases. *Protein Sci.* 13, 854–861.
- Bui, N.K., Gray, J., Schwarz, H., Schumann, P., Blanot, D., and Vollmer, W. (2009). The peptidoglycan sacculus of *Myxococcus xanthus* has unusual structural features and is degraded during glycerol-induced myxospore development. *J. Bacteriol.* 191, 494–505.
- Bui, N.K., Eberhardt, A., Vollmer, D., Kern, T., Bougault, C., Tomasz, A., Simorre, J.P., and Vollmer, W. (2012). Isolation and analysis of cell wall components from *Streptococcus pneumoniae*. *Anal. Biochem.* 421, 657–666.
- Bussiere, D.E., Pratt, S.D., Katz, L., Severin, J.M., Holzman, T., and Park, C.H. (1998). The structure of VanX reveals a novel amino-dipeptidase involved in mediating transposon-based vancomycin resistance. *Mol. Cell* 2, 75–84.
- Cho, J.H., Kim, D.H., Chung, S.J., Ha, N.C., Oh, B.H., and Yong Choi, K. (2002). Insight into the stereochemistry in the inhibition of carboxypeptidase A with N-(hydroxyaminocarbonyl)phenylalanine: binding modes of an enantiomeric pair of the inhibitor to carboxypeptidase A. *Bioorg. Med. Chem.* 10, 2015–2022.
- Christianson, D.W., and Lipscomb, W.N. (1989). Carboxypeptidase A. *Acc. Chem. Res.* 22, 62–69.
- Courtin, P., Miranda, G., Guillot, A., Wessner, F., Mézange, C., Domakova, E., Kulakauskas, S., and Chapot-Chartier, M.P. (2006). Peptidoglycan structure analysis of *Lactococcus lactis* reveals the presence of an L,D-carboxypeptidase involved in peptidoglycan maturation. *J. Bacteriol.* 188, 5293–5298.
- Don, R.H., Cox, P.T., Wainwright, B.J., Baker, K., and Mattick, J.S. (1991). 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Res.* 19, 4008.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501.
- Eschenfeldt, W.H., Lucy, S., Millard, C.S., Joachimiak, A., and Mark, I.D. (2009). A family of LIC vectors for high-throughput cloning and purification of proteins. *Methods Mol. Biol.* 498, 105–115.
- Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* 62, 72–82.
- Fridrich, E., Biboy, J., Adams, C., Lee, J., Ellermeier, J., Gielda, L.D., Dirita, V.J., Girardin, S.E., Vollmer, W., and Gaynor, E.C. (2012). Peptidoglycan-modifying enzyme Pgp1 is required for helical cell shape and pathogenicity traits in *Campylobacter jejuni*. *PLoS Pathog.* 8, e1002602.
- Fridrich, E., Vermeulen, J., Biboy, J., Soares, F., Taveime, M.E., Johnson, J.G., DiRita, V.J., Girardin, S.E., Vollmer, W., and Gaynor, E.C. (2014). Peptidoglycan LD-carboxypeptidase Pgp2 influences *Campylobacter jejuni* helical cell shape and pathogenic properties and provides the substrate for the DL-carboxypeptidase Pgp1. *J. Biol. Chem.* 289, 8007–8018.
- Garcia-Bustos, J.F., Chait, B.T., and Tomasz, A. (1988). Altered peptidoglycan structure in a pneumococcal transformant resistant to penicillin. *J. Bacteriol.* 170, 2143–2147.
- Gorrec, F. (2009). The MORPHEUS protein crystallization screen. *J. Appl. Crystallogr.* 42, 1035–1042.
- Holm, L., and Rosenström, P. (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Res.* 38 (Web Server issue), W545–W549.
- Home, D.S., and Tomasz, A. (1993). Possible role of a choline-containing teichoic acid in the maintenance of normal cell shape and physiology in *Streptococcus oralis*. *J. Bacteriol.* 175, 1717–1722.
- Kabsch, W. (2010). XDS. *Acta Crystallogr. D Biol. Crystallogr.* 66, 125–132.
- Korndörfer, I.P., Kanitz, A., Danzer, J., Zimmer, M., Loessner, M.J., and Skerra, A. (2008). Structural analysis of the L-alanoyl-D-glutamate endopeptidase domain of *Listeria bacteriophage* endolysin Ply500 reveals a new member of the LAS peptidase family. *Acta Crystallogr. D Biol. Crystallogr.* 64, 644–650.
- Korza, H.J., and Bochtler, M. (2005). *Pseudomonas aeruginosa* LD-carboxypeptidase, a serine peptidase with a Ser-His-Glu triad and a nucleophilic elbow. *J. Biol. Chem.* 280, 40802–40812.
- Koshland, D.E. (1958). Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl. Acad. Sci. USA* 44, 98–104.
- Lovering, A.L., Safadi, S.S., and Strynadka, N.C.J. (2012). Structural perspective of peptidoglycan biosynthesis and assembly. *Annu. Rev. Biochem.* 81, 451–478.
- Marcyjaniak, M., Odintsov, S.G., Sabala, I., and Bochtler, M. (2004). Peptidoglycan amidase MepA is a LAS metallopeptidase. *J. Biol. Chem.* 279, 43982–43989.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* 40, 658–674.

- Metz, R., Henning, S., and Hammes, W.P. (1986). LD-carboxypeptidase activity in *Escherichia coli*. II. Isolation, purification and characterization of the enzyme from *E. coli* K 12. *Arch. Microbiol.* **144**, 181–186.
- Meziane-Cherif, D., Stogios, P.J., Evdokimova, E., Savchenko, A., and Courvalin, P. (2014). Structural basis for the evolution of vancomycin resistance D,D-peptidases. *Proc. Natl. Acad. Sci. USA* **111**, 5872–5877.
- Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006). HKL-3000: the integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta Crystallogr. D* **62**, 859–866.
- Pape, T., and Schneider, T.R. (2004). HKL2MAP: a graphical user interface for phasing with SHELX programs. *J. Appl. Crystallogr.* **37**, 843–844.
- Royet, J., Gupta, D., and Dziarski, R. (2011). Peptidoglycan recognition proteins: modulators of the microbiome and inflammation. *Nat. Rev. Immunol.* **11**, 837–851.
- Sauvage, E., Kerff, F., Terrak, M., Ayala, J.A., and Charlier, P. (2008). The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS Microbiol. Rev.* **32**, 234–258.
- Sheldrick, G.M. (2008). A short history of SHELX. *Acta Crystallogr. A* **64**, 112–122.
- Sycuro, L.K., Pincus, Z., Gutierrez, K.D., Biboy, J., Stern, C.A., Vollmer, W., and Salama, N.R. (2010). Peptidoglycan crosslinking relaxation promotes *Helicobacter pylori*'s helical shape and stomach colonization. *Cell* **141**, 822–833.
- Sycuro, L.K., Wyckoff, T.J., Biboy, J., Born, P., Pincus, Z., Vollmer, W., and Salama, N.R. (2012). Multiple peptidoglycan modification networks modulate *Helicobacter pylori*'s cell shape, motility, and colonization potential. *PLoS Pathog.* **8**, e1002603.
- Sycuro, L.K., Rule, C.S., Petersen, T.W., Wyckoff, T.J., Sessler, T., Nagarkar, D.B., Khalid, F., Pincus, Z., Biboy, J., Vollmer, W., and Salama, N.R. (2013). Flow cytometry-based enrichment for cell shape mutants identifies multiple genes that influence *Helicobacter pylori* morphology. *Mol. Microbiol.* **90**, 869–883.
- Templin, M.F., Ursinus, A., and Höltje, J.V. (1999). A defect in cell wall recycling triggers autolysis during the stationary growth phase of *Escherichia coli*. *EMBO J.* **18**, 4108–4117.
- Vollmer, W., and Tomasz, A. (2001). Identification of the teichoic acid phosphorylcholine esterase in *Streptococcus pneumoniae*. *Mol. Microbiol.* **39**, 1610–1622.
- Vollmer, W., Blanot, D., and de Pedro, M.A. (2008a). Peptidoglycan structure and architecture. *FEMS Microbiol. Rev.* **32**, 149–167.
- Vollmer, W., Joris, B., Charlier, P., and Foster, S. (2008b). Bacterial peptidoglycan (murein) hydrolases. *FEMS Microbiol. Rev.* **32**, 259–286.
- Young, F.E., and Spizizen, J. (1961). Physiological and genetic factors affecting transformation of *Bacillus subtilis*. *J. Bacteriol.* **81**, 823–829.

Attachment of Capsular Polysaccharide to the Cell Wall in *Streptococcus pneumoniae*

Alice Eberhardt,^{1,*†} Christopher N. Hoyland,^{1,*} Daniela Vollmer,^{1,*} Stephanie Bisle,^{1,‡}
Robert M. Cleverley,¹ Ola Johnsborg,^{2,§} Leiv S. Håvarstein,² Richard J. Lewis,¹ and Waldemar Vollmer¹

Streptococcus pneumoniae protects itself from components of the human immune defense system by a thick polysaccharide capsule, which in most serotypes is covalently attached to the cell wall peptidoglycan. Members of the LytR-Cps2A-Psr (LCP) protein family have recently been implicated in the attachment of anionic polymers to peptidoglycan in Gram-positive bacteria, based on genetic evidence from *Bacillus subtilis* mutant strains and on the crystal structure of *S. pneumoniae* Cps2A containing a tightly bound polyprenol (pyro)phosphate lipid. Here, we provide evidence that Cps2A and its two pneumococcal homologs, LytR and Psr, contribute to the maintenance of normal capsule levels and to the retention of the capsular polysaccharide at the cell wall in the capsular type 2 *S. pneumoniae* strain D39. GFP fusions of all three LCP proteins showed enhanced localization at mid-cell, indicating a role in cell wall growth. Single *cps2A* or *psr* mutants produced a reduced amount of capsule. A *cps2A lytR* double mutant showed greatly impaired growth and cell morphology and lost approximately half of the total capsule material into the culture supernatant. We also present the crystal structure of the *B. subtilis* LCP protein YwtF and provide crystallographic evidence for the phosphotransferase activity of Cps2A, supporting an enzymatic function in the attachment of capsular polysaccharides to cell wall peptidoglycan.

Introduction

STREPTOCOCCUS PNEUMONIAE IS A MAJOR human pathogen capable of causing otitis media and acute sinusitis but also severe invasive diseases, such as pneumoniae, meningitis, bacteremia, and sepsis. Mainly young children, the elderly, and immune-compromised patients are at risk, and each year ~1 to 2 million patients die from pneumococcal infections.⁴³ The cell surface of the Gram-positive bacterium *S. pneumoniae* is made of a peptidoglycan-teichoic acid cell wall^{47,48} covered and shielded by a polysaccharide capsule, allowing it to evade recognition by the innate immune system of the host.^{6,35} Cell-wall-anchored proteins^{16,18} and structural modifications in peptidoglycan, such as N-deacetylation^{45,46} and O-acetylation,^{11,44} contribute to pneumococcal virulence. The most relevant virulence determinant is the capsule, and nonencapsulated pneumococcal mutants are avirulent. To date, 93 chemically distinct capsule types have been described.^{7,9,20,37} Most capsular types are

tightly associated with the bacterial cell wall. From 12 capsular polysaccharides tested, all but one (serotype 3) were covalently linked to the cell wall, most likely via phosphodiester bonds to the N-acetylmuramic acid residues in the glycan chains of the peptidoglycan.³⁸ The peptidoglycan also carries a complex wall teichoic acid, an anionic cell wall polymer that binds cell wall proteins and divalent cations,^{18,3,25} the biosynthesis of which is the topic of a companion article.¹² Novel methods to fractionate and analyze the complex structure of pneumococcal cell wall have recently been published.⁸

In the majority of serotypes, the capsular polysaccharide genes are located between *dexB* and *aliA*.⁵ Here, we focus on the serotype 2 strain, D39, the capsular region of which begins with the *cps2A–D* genes (Fig. 1A). All 17 capsular genes in this region are under control of the promoter upstream of *cps2A*.¹⁷ The first gene in the region, *cps2A*, encodes for a member of the LCP protein family (LytR, Cps2A, and Psr homologs), which is widespread in Gram-positive bacteria

¹Centre for Bacterial Cell Biology, Institute for Cell and Molecular Biosciences, Newcastle University, Newcastle upon Tyne, United Kingdom.

²Department of Chemistry, Biotechnology, and Food Science, Norwegian University of Life Sciences, Ås, Norway.

*These three authors contributed equally to this work.

[†]Present address: Department of Microbiology, Karolinska Institutet, Tumor and Cell Biology, Stockholm, Sweden.

[‡]Present address: Mikrobiologisches Institut—Klinische Mikrobiologie, Immunologie und Hygiene, Universitätsklinikum Erlangen, Erlangen, Germany.

[§]Present address: Nofima, Ås, Norway.

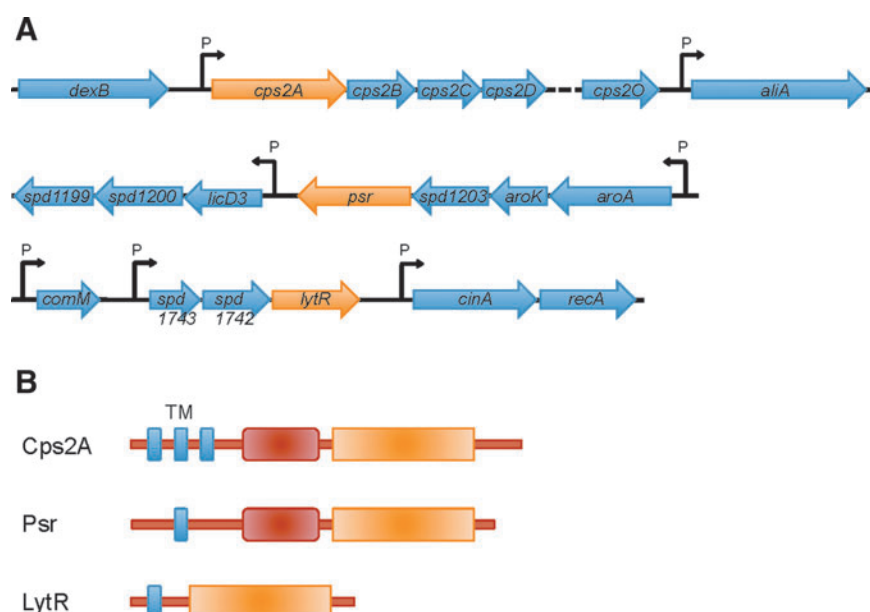


FIG. 1. The LCP genes and proteins in *Streptococcus pneumoniae*. **(A)** Gene regions of *cps2A*, *psr*, and *lytR* in strain D39. *Cps2A* is the first gene of the capsule region and followed by 16 other genes involved in capsular biosynthesis and assembly. *Psr* is the last gene in an operon with *aroA*, *aroK*, and *spd1203*. *LytR* is the last gene in an operon with two genes of unknown function located upstream. **(B)** Organization of domains in Cps2A, Psr, and LytR. Cps2A contains 3 predicted transmembrane domains, whereas Psr and LytR possess 1 (indicated in blue). The LCP domains are shown in light red; an additional domain present in Cps2A and Psr is shown in dark red.

and does not occur in Gram negatives. The pleiotropic phenotypes of LCP mutant strains of various species have led to the suggestion that this protein family is a transcriptional regulator of cell wall processes.²⁷ Cps2A (also called Wzg), Cps2B (Wzh), Cps2C (Wcd), and Cps2D (Wze) have all been implicated in the regulation of capsule synthesis in *S. pneumoniae*. Indeed, Cps2D is an autophosphorylating tyrosine kinase that appears to be regulated by Cps2C^{4,30} and both might be spatial regulators that connect capsule synthesis to cell division.²¹ Cps2A is required for the full expression of capsule; *cps2A* mutant strains contain significantly less capsule material attached to the cell wall.³¹ However, like other LCP family members, Cps2A is an integral membrane protein, the bulk of which is located on the outside face of the cytoplasmic membrane (Fig. 1B)¹⁹; it is not immediately apparent that how to reconcile this observation with any role in DNA binding. A recent report by Kawai *et al.* provides strong evidence that LCP proteins are the long searched for enzymes catalyze the covalent attachment of anionic cell wall polymers, like teichoic acids and capsular polysaccharides, to peptidoglycan.²⁴ This publication detailed genetic and biochemical evidence from the characterization of mutant strains of *Bacillus subtilis* and the crystal structure of the soluble part of Cps2A. Interestingly, the conserved, LCP domain of Cps2A contained a (pyro)phosphorylated polyprenol lipid, the precise chemical identity of which was dependent upon the *Escherichia coli* expression strain used for purification of the protein. Preliminary assays showed that Cps2A had weak, metal-ion-dependent pyrophosphatase activity, supporting an enzymatic function. Most likely, Cps2A binds the undecaprenol-pyrophosphoryl-linked capsule precursor to transfer the phosphorylated capsule chain to C6-OH of MurNAc residues in the peptidoglycan precursor or polymer.

In this study we have investigated the role of Cps2A, LytR, and Psr in the attachment of capsular polysaccharides to the cell wall of capsule type 2 strain D39. We show that all LCP proteins localize to the cell membrane and are enriched at mid-cell. *Cps2A* and *psr* mutants have a reduced amount of capsule. Inactivation of *lytR* proved difficult in the background of the encapsulated D39 strain, leading to mutants that were not viable in liquid culture. A suppressor mutant of a D39 *cps2A-lytR* double mutant grew with strongly distorted morphology in liquid culture, where it lost a large portion of its capsule material into the supernatant. LytR, together with Cps2A, thus plays a role in retaining the capsule at the cell wall. The crystal structure of the LytR homolog from *B. subtilis*, YwtF, and the pyrophosphorylase observed by Cps2A support an enzymatic function of these proteins. These data and our previous observations²⁴ are consistent with a model whereby Cps2A is responsible for the covalent attachment of capsular polysaccharide to the pneumococcal cell wall, and that LytR can take over this function in the absence of Cps2A.

Materials and Methods

Bacterial strains, growth conditions, and genetic transformation

Table 1 contains the strains used in this study. *S. pneumoniae* R6^{26,36} and D39² were grown at 30°C or 37°C in C+Y (pneumococcal C medium with yeast extract) medium containing 1 mg/ml yeast extract.⁴¹ To induce the expression of *gfp+* fusions under *P_{cpsD}* control on plasmid pJWV25 and its derivatives, 0.15 mM ZnCl₂ was added to the C+Y medium. Blood agar plates were made of tryptic soy broth (TSB) medium containing 1.5% agar and 3% defibrinated sheep

TABLE 1. STRAINS USED IN THIS STUDY

Strain	Properties	Reference or source
D39	clinical isolate, serotype 2	2
R6	unencapsulated, nonvirulent descendent of D39	26,36
R6 <i>P_{czcD}-gfp</i> + - <i>cps2A</i>	<i>tetM</i> , <i>P_{czcD}-gfp</i> + - <i>cps2A</i>	This study
D39 <i>P_{czcD}-gfp</i> + - <i>cps2A</i>	<i>tetM</i> , <i>P_{czcD}-gfp</i> + - <i>cps2A</i>	This study
R6 <i>P_{czcD}-gfp</i> + - <i>lytR</i>	<i>tetM</i> , <i>P_{czcD}-gfp</i> + - <i>lytR</i>	This study
D39 <i>P_{czcD}-gfp</i> + - <i>lytR</i>	<i>tetM</i> , <i>P_{czcD}-gfp</i> + - <i>lytR</i>	This study
R6 <i>P_{czcD}-gfp</i> + - <i>psr</i>	<i>tetM</i> , <i>P_{czcD}-gfp</i> + - <i>psr</i>	This study
D39 <i>P_{czcD}-gfp</i> + - <i>psr</i>	<i>tetM</i> , <i>P_{czcD}-gfp</i> + - <i>psr</i>	This study
D39 <i>cps2A::ermB</i>	<i>ermB</i> , <i>cps2A::ermB</i>	This study
D39 <i>cps2A::ermB P_{czcD}-gfp</i> + - <i>cps2A</i>	<i>tetM</i> , <i>ermB</i> , <i>cps2A::ermB</i> , <i>P_{czcD}-gfp</i> + - <i>cps2A</i>	This study
D39 <i>psr</i>	<i>ermB</i> , insertion-inactivation of <i>psr</i> using pJDC9	This study
D39 <i>cps2A::ermB lytR::kan P_{czcD}-lytR</i>	<i>cps2A::ermB</i> , <i>lytR::kan</i> , <i>P_{czcD}-lytR</i>	This study
D39 <i>cps2A::ermB lytR::kan</i>	<i>cps2A::ermB</i> , <i>lytR::kan</i>	This study
D39 <i>cps2A::ermB psr::kan</i>	<i>cps2A::ermB</i> , <i>psr::kan</i>	This study

blood. For growth curves, *S. pneumoniae* was grown overnight at 37°C on TSB blood agar plates with the appropriate antibiotics. C + Y medium without antibiotics was inoculated from colonies grown on plates to an OD₆₂₀ of 0.05–0.07, incubated in a water bath at 37°C, and the OD₆₂₀ was measured at the indicated time intervals.

E. coli DH5α (Invitrogen) was grown in lysogeny broth (LB) medium at 37°C with aeration. Where necessary, erythromycin was added at a concentration of 1 mg/ml for *E. coli* and 1 µg/ml for *S. pneumoniae*, tetracycline was added at 12.5 µg/ml for *E. coli* and 2.5 µg/ml for *S. pneumoniae*, and kanamycin was added at 50 µg/ml for *E. coli* and 400 µg/ml for *S. pneumoniae*.

Transformation of *S. pneumoniae*

DNA was introduced to the pneumococcal strains by means of natural transformation using 100 ng/ml of CSP-1 to induce competence in the recipient strain.³² Cultures of *S. pneumoniae* were grown at 37°C until they reached an OD₆₂₀ of 0.15–0.2, at which point CSP-1 and the transforming DNA were added, and the cultures were further incubated for 90–120 min at 37°C before plating on selective medium to identify transformants. *S. pneumoniae* transformants were selected by plating on TSB blood agar plates supplemented with the appropriate antibiotic.

Polymerase chain reaction amplification, cloning, and sequencing

Chromosomal DNA from *S. pneumoniae* was isolated using the Promega Wizard® genomic DNA purification system. For this, pneumococcal cells from liquid culture were harvested by centrifugation and resuspended in lysis buffer (50 mM Tris/HCl [pH 8.5], 5 mM MgCl₂, 25 µg/ml RNase, and 0.4% DOC). To obtain chromosomal DNA, single colonies were propagated on TSB blood plates after transformation and cell material from the plate was resuspended in lysis buffer, and then incubated at 37°C for 10 min. Plasmids were purified with the Qiagen plasmid midi kit. Restriction enzymes and T4 ligase were purchased from Fermentas. DNA fragments were purified by using the QIAquick® PCR purification kit. DNA was amplified with Phusion® high-fidelity DNA polymerase from Finnzymes. Oligonucleotides listed in Table 2 were supplied by Sigma. Sequencing was performed at GATC, Germany.

Cloning of *gfp* + fusions

For production of GFP + -tagged proteins in *S. pneumoniae*, the zinc-inducible promoter of plasmid pJWV25 was used.¹³ To generate *gfp* + fusions of *cps2A*, *lytR*, and *psr* in the background of R6 and D39, *cps2A* was amplified using the oligonucleotides *cps2A_F* and *cps2A_R*, *lytR* was amplified using the oligonucleotides *lytR_F* and *lytR_R*, and *psr* was amplified using the oligonucleotides *psr_F* and *psr_R*. The polymerase chain reaction (PCR) products were digested with the appropriate restriction enzyme (Table 2), ligated with digested pJWV25, and transformed into *E. coli* DH5α selecting on tetracycline. Transformants were screened by colony PCR using oligonucleotides pJWV25_scr_F and pJWV25_scr_R. The integrity of the insert in pJWV25 was confirmed by sequencing using oligonucleotides pJWV25_scr_F and pJWV25_scr_R. Correct plasmids were transformed into R6 and D39 while selecting for tetracycline resistance, resulting in strains R6 *P_{czcD}-gfp* + -*cps2A*, D39 *P_{czcD}-gfp* + -*cps2A*, R6 *P_{czcD}-gfp* + -*lytR*, D39 *P_{czcD}-gfp* + -*lytR*, R6 *P_{czcD}-gfp* + -*psr*, and D39 *P_{czcD}-gfp* + -*psr*. The correct integration of the pJWV25 derivatives into the *bgaA* region was confirmed by PCR amplification with the oligonucleotides *bgaA_check_F* and *bgaA_check_R*.

Gene deletions

D39 *cps2A::ermB* was created by replacing the *cps2A* gene by *ermB* followed by the promoter of *cps2A* to ensure native transcription levels of the downstream genes. Therefore, a construct consisting of four PCR fragments was consecutively cloned into pGEM-3Z: the 1 kb upstream flanking region of *cps2A* (oligonucleotides *cps2A_up_F* and *cps2A_up_R*), the *ermB* gene (oligonucleotides *ermB_F* and *ermB_R*), the promoter region of *cps2A* (oligonucleotides *Pcps2A_F* and *Pcps2A_R*), and the 1 kb downstream flanking region of *cps2A* (oligonucleotides *cps2A_down_F* and *cps2A_down_R*). Transformants of DH5α were screened by colony PCR using the oligonucleotides pGEM3Z_scr_F and pGEM3Z_scr_R. The final knockout construct was sequence verified and transformed into D39, selecting for erythromycin resistance. To confirm correct integration of the construct into the capsular region, the locus was PCR amplified with the oligonucleotides *Δcps2A_seq_F* and *Δcps2A_seq_R* and sequenced. D39 *cps2A::ermB* was transformed with pJWV25-*cps2A*, resulting

TABLE 2. OLIGONUCLEOTIDES USED IN THIS STUDY

Name	Sequence	Restriction site
cps2A_F	CAGCAGGCTAGCAGTAGACGTTTTAAAAAATCACGTTACACAG	NheI
cps2A_R	CATCATGCGGCCGCTCATCTACCCTCCATCACATCCTGTATAGC	NotI
lytR_F	CAGCAGACTAGTGTTAAAAAAATTATTGGAATGGTGCTAGC	SpeI
lytR_R	CATCATGCGGCCGCTTAATTATCTTCATCACCAACAGGTGC	NotI
psr_F	CAGCAGACTAGTAGTAAAGAAAAATCCCCTTAAGTC	SpeI
psr_R	CATCATGCGGCCGCTTATCTCAAGCCCATTGAGCG	NotI
pJWV25_scr_F	GACAACCATTACCTGTCGAC	None
pJWV25_scr_R	TTGGTGCAAGGAAGGTCATC	None
bgaA_check_F	CCACTCGCAACAATCACTTGG	None
bgaA_check_R	GGTTGAGAACTATTTTGTTC	None
cps2A_up_F	GACGACGAATTCCAAATTCCTTACTTAAAGG	EcoRI
cps2A_up_R	GACGATGGTACCGTTCTCTCTTATTGTGAGCG	KpnI
ermB_F	GACGACGGTACCACAAGTGATTTGTGATTGTTG	KpnI
ermB_R	GACGACGGATCCTTATTTCTCCCGTTAAATAATAG	BamHI
Pcps2A_F	GACGACGGATCCTGTGTACTATTATATTGAAAC	BamHI
Pcps2A_R	GACTAGGTCGACGATTAACACATTATACATTG	Sall
cps2A_down_F	TATACGTCGACGATGGAGGGTAGATGAAATG	Sall
cps2A_down_R	TACTACAAGCTTCGAAAGGATAATCTCACGG	HindIII
pGEM3Z_scr_F	AGTCACGACGTTGTAAACG	None
pGEM3Z_scr_R	AGCGCAACGCAATTAAGTTG	None
psr_F	CATCATGAATTCTCCCTTAAGTCATCATGAG	EcoRI
psr_R	CATCATCCTGCGAGGATTTCTTATCTGAGCCCC	SbfI
intR_pJDC9	TTGTGTGGAATTGTGAGCGG	None
intF_pJDC9	TTCGCTATTACGCCAGCTGG	None
intF_psr	ACTCAAGACAGCACTGGGTG	None
intR_psr	TTTCTGCCCTTTTGTCCCC	None
Δ cps2A_seq_F	GCAAGACTTGACAGTAGAAGG	None
Δ cps2A_seq_R	CGAACATTTGGAGAAGAAGG	None
lytR_ext_F	AGACTGCTACGTTGACTGGC	None
lytR_ext_R	TCGTAATCAAGCAGCAGTCC	None
lytR_int_F	TATAGCCTAGCAAGTTAGGG	None
lytR_int_R	GAATCAGGGAATGGTCAGGC	None
bgaA_int_F	CATCCGTACAACCCACAACC	None
bgaA_int_R	ATTACGCTCAGGTCCATTGC	None
seq_psr_middle_F	CCAAGATGCTAGCGCAAAGA	None
psr_flank_R	ATCTCCAGACTACGAGCCTG	None
YwtF FW	GCTTTTGTAAGCTGACCATGGCTTCTAAGGAAGC	NcoI
YwtF RVII	GGCGTAATCGCCATCCCGTCTCTTATACACCTAAATCG	None
YwtF RV	GGCGTAATCGCCATCCCGTCTCGAGTACACCTAAATCG	XhoI

in D39 *cps2A::ermB* *P_{czcD}-gfp* + *-cps2A*, which was complemented with *gfp* + *-cps2A* upon zinc induction.

D39*psr* contains an insertion-inactivation of *psr*. This mutant was generated using the pJDC9-based insertion-deletion mutagenesis method.¹⁰ The first 500bp of *psr* (omitting the start codon ATG) was cloned in pJDC9 using the oligonucleotides *psr_F* and *psr_R*. The insert was verified by sequencing and the plasmid was transformed into D39. Correct integration of the plasmid into the *psr* gene was verified by PCR amplification. With oligonucleotides *intF_psr* located in the chromosome and *intR_pJDC9* on the plasmid, the upstream region of *psr* was amplified. The downstream region was checked by PCR using oligonucleotides *intF_pJDC9* and *intR_psr*.

Construction of D39 *cps2A::ermB* *lytR::kan* *P_{czcD}-lytR* and D39 *cps2A::ermB* *psr::kan*

To construct the D39 *cps2A::ermB* *lytR::kan* mutant, a kanamycin resistance cassette flanked by DNA corresponding to the upstream and downstream regions of the *psr* locus was designed. A DNA fragment corresponding to the up-

stream region of *psr* was amplified by PCR from strain D39 and subsequently fused to the 5'-end of the kanamycin cassette in a second PCR. Next, a PCR fragment corresponding to the downstream region of *psr* was amplified and fused to the 3'-end of the resistance cassette. The final fragment was integrated into the *psr* locus of D39 *cps2A::ermB* by natural transformation.

To construct the D39 *cps2A::ermB* *lytR::kan* *P_{czcD}-lytR* strain, pJWV25 was inserted into the genome of the D39 *cps2A::ermB* strain as described previously.¹³ The resulting strain was subsequently made streptomycin resistant by transformation with genomic DNA from strain CP1415.³³ Next, a *kan::rpsL* Janus cassette³⁹ was integrated downstream of the zinc-inducible *P_{czcD}* promoter of the chromosomal pJWV25. The Janus cassette was subsequently replaced by a PCR fragment consisting of the coding sequence of D39 *lytR*. Finally, the wt⁺ copy of *lytR* was removed by natural transformation with a PCR cassette as described previously,²² and colonies were selected on agar containing 0.15 mM ZnCl₂.

D39 *cps2A::ermB* *lytR::kan* was created by transforming D39 *cps2A::ermB* with chromosomal DNA isolated from D39

cps2A::ermB lytR::kan P_{czcD}-lytR and selecting colonies for kanamycin resistance and, therefore, the replacement of *lytR* with the kanamycin resistance cassette. The resultant colonies were screened by PCR for the absence of *lytR* at its locus using the oligonucleotides *lytR_ext_F* and *lytR_ext_R*. The general absence of *lytR* was verified using *lytR_int_F* and *lytR_int_R* and the absence of insertions at *bgaA* locus was checked using *bgaA_int_F* and *bgaA_int_R*. To exclude the possibility that any mutation occurred in the *psr* gene, *psr* was amplified and sequenced using *intF_psr*, *seq_psr_midle_F*, and *psr_flank_R*.

Expression of GFP+ fusion proteins in S. pneumoniae and fluorescence microscopy

Strains expressing GFP+ fusion proteins were grown and samples were prepared as published previously.¹³ The GFP+ signal was recorded using a DeltaVision experimental setup and the application of an exposure time of 2 sec using an excitation filter (Chroma) and 100% transmission (excitation at 490 nm and emission at 535 nm). For phase-contrast microscopy, the specimens were exposed for 100 ms with 100% transmission. Images taken were deconvoluted using the DeltaVision deconvolution software (ratio conservative, number of cycles 15, wavelength 528 nm, medium noise filtering).

Western blot and dot blot analyses

The expression and the correct size of the GFP+ fusion proteins were verified by western blotting. Whole cell lysates of the strains expressing GFP+ fusions were run in equal amounts on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The proteins were blotted onto nitrocellulose and the GFP+ fusions were detected with anti-GFP antibody (rabbit polyclonal; Invitrogen). A secondary horseradish peroxidase (HRP)-labeled anti-rabbit antibody (goat polyclonal; Invitrogen) was used in combination with an ECL kit (Amersham) for visualization of the antibody complexes.

Quantification of capsule per colony

D39 *cps2A::ermB*, D39, and R6 were transformed with genomic DNA from D39 *cps2A::ermB lytR::kan P_{czcD}-lytR* and selected on TSB blood plates containing kanamycin. Agar pieces containing a single colony were excised from the plates and stored at -20°C . Thawed samples were mixed with 60 μl of lysis buffer and the agar was melted for 10 min at 100°C . In total, 7 serial 1:3 dilution steps in H_2O were made and 2 μl of each dilution was applied to a nitrocellulose membrane. The membrane was air-dried for 2 h at room temperature and 15 min at 60°C , blocked in TBS + 0.5% casein for 1 h at room temperature, incubated with anti-serotype 2 serum in 1:10,000 (from rabbit; Statens Serum Institute) in TBS + 0.5% casein overnight at 4°C , washed 3 times for 5 min in TBS + 0.2% Tween20, incubated with a secondary HRP-labeled anti-rabbit antibody (goat polyclonal; Invitrogen), washed 3 times for 5 min in TBS + 0.2% Tween20, and detected using an ECL kit (NBS Biologicals) on photographic films (GE Healthcare).

Quantification of capsule in culture lysate and supernatant

D39 *cps2A::ermB*, D39*psr*, D39 *cps2A::ermB lytR::kan*, D39 *cps2A::ermB psr::kan*, D39, and R6 were grown on plates

containing the appropriate antibiotic. C+Y medium was inoculated from plates to OD₆₂₀ of 0.08 and grown at 37°C . To quantify total capsular material, an aliquot of culture was lysed at OD₆₂₀ 0.2 with 0.4% DOC, DNase, and RNase (10 $\mu\text{g}/\text{ml}$ each) at 37°C for 10 min and centrifuged at 14,000 rpm for 10 min at room temperature. The supernatant was stored at -20°C . A second aliquot of the culture was centrifuged at 13,000 rpm for 10 min at 4°C , and the supernatant was mixed with 0.4% sodium deoxycholate before storage at -20°C . Samples were thawed at 50°C and diluted 1:1.5 in 12 steps with H_2O . About 2 μl of each dilution was applied onto nitrocellulose membrane, air-dried, and processed as described for colony blots, but here the samples were washed with TBS without Tween20. The electrochemiluminescence signal was detected using Image Quant LAS 4,000 mini (GE Healthcare).

Visualization of proteins in culture lysate and supernatant

Whole culture, or culture supernatant (after centrifuging at 13,000 rpm for 15 min at 4°C) collected at an OD₆₂₀ of 0.2 was mixed with sample buffer, and boiled 4 times for 10 min at 100°C and vortexed for 1 min. The samples were stored at -20°C before separation by 12% SDS-PAGE and visualization by silver staining (Proteo Silver Staining Kit; Sigma).

Capsule negative staining with Congo Red

The capsule was visualized by mixing 1 μl Congo Red solution (2% Congo Red in H_2O) with 1 μl of liquid culture (or a colony picked from plate) on a glass slide. The mixture was spread on the slide and let to air-dry. Cells were visualized with a Nikon Eclipse Ti microscope with phase contrast using a 100 \times oil-immersion objective. The capsule was visible as a clear area surrounding darker cell bodies on red background.

Electron microscopy

Cells were grown in C+Y at 37°C . At an OD₆₂₀ of 0.23 (strain D39 *cps2A::ermB lytR::kan*) or 0.4 (D39), samples were fixed in 2% glutaraldehyde in sodium cacodylate buffer (pH 7.4) (TAAB Laboratory Equipment) overnight. Samples were processed by the Electron Microscopy Research Service of Newcastle University. Briefly, samples were incubated in 1% osmium tetroxide (Agar Scientific) for 1 h, and then dehydrated in an acetone-graded series before being impregnated with a graded series of epoxy resin (TAAB Laboratory Equipment) in acetone and finally embedded in 100% resin and set at 60°C for 24 hr. Cells were sectioned and counterstained with 2% uranyl acetate and lead citrate (Leica) before being imaged on a Philips CM100 Compustage Transmission Electron Microscope (FEI) with an AMT CCD camera (Deben).

Cloning of B. subtilis ywtF and purification and crystallization of YwtF

The genetic region encoding the soluble domain of YwtF (amino acids 44–322, $\Delta\text{tm-YwtF}$) was amplified from *B. subtilis* 168 genomic DNA by PCR, using the primers YwtF FW and YwtF RVII for the first reaction and this PCR product, with the primers YwtF FW and YwtF RV, for the

second round of PCR. The PCR products were digested with the appropriate restriction enzymes (Table 2), ligated with similarly restricted pET28a, and transformed into *E. coli* XL-1 Blue, selecting for kanamycin resistance. The product of this reaction, pET28a- Δ tm-YwtF, placed the open-reading frame for Δ tm-YwtF under a T7, IPTG-inducible promoter and added a C-terminal hexahistidine tag to the expressed protein. Sequencing confirmed the correct DNA sequence of pET28a- Δ tm-YwtF.

Expression and purification of Δ tm-YwtF

pET28a- Δ tm-YwtF was transformed into *E. coli* BL21(DE3) selecting for kanamycin resistance and grown at 37°C with shaking in LB media containing kanamycin to an OD₅₉₅ of 0.6 before expression of Δ tm-YwtF was induced by the addition of 1 mM IPTG. The temperature of the culture was reduced to 18°C and the growth was continued for a further 16 h before the cells were pelleted by centrifugation at 4,000 g at 4°C (Beckman J-26-XP, JLA 8.1000 rotor).

Cell pellets were resuspended in 20 ml of Ni-NTA buffer A (50 mM Tris-HCl [pH 8.0], 300 mM NaCl, and 10 mM imidazole) before the addition of 1 ml of 25× EDTA-free complete protease inhibitor (Roche Diagnostics) and 20 µl of DNase (Sigma; final concentration of 5 µg/ml). The cell suspension was sonicated (Bandelin sonopuls HD2070) on ice for 3 min at 80% power and 70% cycle before being clarified by centrifugation at 20,000 g, 4°C for 25 min (Beckman J-26-XP with JA 25.50 rotor).

Soluble cell lysate was loaded manually using an ÄKTA prime (GE Healthcare) at 1 ml/min onto a 5 ml Ni²⁺-NTA (Ni-NTA) cartridge (Qiagen), pre-equilibrated in Ni-NTA buffer A. The column was washed with 40 ml of Ni-NTA buffer A to remove any nonspecifically bound proteins. The tightly bound proteins were eluted with Ni-NTA buffer B (50 mM Tris-HCl [pH 8.0], 300 mM NaCl, and 250 mM imidazole). The presence of protein was monitored by absorption at 280 nm and confirmed by SDS-PAGE of the collected fractions.

Fractions containing the protein of interest were pooled and concentrated (Amicon Ultra-15 10K; Millipore) to 1 ml before size-exclusion chromatography on a Superdex S75 HiLoad 16/60 (GE healthcare) column pre-equilibrated in gel filtration buffer (10 mM Tris-HCl [pH 8.0] and 250 mM NaCl) using an ÄKTA purifier (GE Healthcare). Eluted proteins were fractionated by volume and monitored by absorption at 280 nm. SDS-PAGE confirmed the presence and purity of Δ tm-YwtF.

Crystallization

Crystals of Δ tm-YwtF were obtained in hanging drops by mixing equal volumes of protein at 25 mg/ml in 10 mM Tris (pH 7.5), 100 mM NaCl with a well solution of 0.1 M HEPES (pH 7.5), 20% PEG-3350, 0.2 M MgCl₂ at 22°C. Crystals were cryoprotected by first transferring them into a drop of the well solution, followed by a drop of well solution supplemented with 10% PEG 400, before flash freezing in liquid nitrogen.

Data collection and processing

Diffraction data were collected at the Diamond Light Source synchrotron on beamline I02. A total of 500 images of

0.4° oscillations were collected at a wavelength of 0.9795 Å for Δ tm-YwtF. Diffraction data were integrated using MOSFLM²⁸ and scaled and reduced with SCALA.¹⁵ The crystal structure was solved by molecular replacement using MolRep⁴² from the CCP4 suite (Collaborative Computational Project, Number 4, 1994) and with the previously solved, lower-resolution crystal structure of Δ tm-YwtF (unpublished; PDBid 3MEJ) as the search model. The correctly positioned protein model was subjected to one round of rigid body refinement in REFMAC5,³⁴ to optimize its location in the crystallographic asymmetric unit. The resultant protein model was subjected to numerous cycles of restrained refinement in PHENIX REFIN¹ interspersed with manual model building in COOT¹⁴ until the refinement converged (Table 3).

Mass spectrometry of Δ tm-YwtF

The purified protein was dialyzed extensively against 10 mM ammonium acetate (pH 8.0) and then concentrated to 1 mg/ml. Samples were analyzed at the Mass Spectrometry facility at the Astbury Centre for Structural Molecular Biology at the University of Leeds, United Kingdom.

Purification and crystallization of Cps2A(R267A)

Cps2A(R267A) was purified by Ni-NTA affinity and size-exclusion chromatography as described previously.²⁴ Crystals were grown in hanging drops with well solutions of either 0.1 M sodium acetate (pH 4.6), 8% PEG-4K or 0.2 M di-ammonium citrate (pH 5.0), 15% PEG-3350. The former crystals were cryoprotected by direct transfer to 0.1 M sodium acetate (pH 4.6), 6% PEG-4K, 30% ethylene glycol and then flash frozen in liquid nitrogen, whereas the latter were transferred to 0.2 M di-ammonium citrate (pH 5.0), 15% PEG-3350, 20% PEG 300 before flash freezing in liquid nitrogen.

Diffraction data were integrated using MOSFLM²⁸ for the crystals grown in acetate and with XDS²³ for the crystals grown in citrate; both datasets were scaled with SCALA.¹⁵ An initial model was obtained by fitting the structure of the wild-type protein (PDB entry 2XXP²⁴) into the data by molecular replacement with PHASER.²⁹ The structures were subsequently refined by iterative rounds of refinement (with REFMAC³⁴ for acetate crystals and PHENIX¹ for citrate crystals) and rebuilding in COOT.¹⁴

Phosphate release assay

Fifty-microliter samples of Cps2A and Cps2A(R267A) (8 mg/ml) were incubated for 5 days at 22°C in 20 mM di-ammonium citrate (pH 5.0). About 6 µl of a stock solution of 1 M Tris-HCl (pH 8.0) was added and the release of inorganic phosphate was measured using the colorimetric PiPer assay (Invitrogen).

Results

Based on the role of LCP proteins in the cell wall assembly process of *B. subtilis*, and the crystal structure of the soluble part of Cps2A, we inferred that the three homologs in *S. pneumoniae* (Cps2A, LytR, and Psr; Fig. 1) might have roles in the attachment of capsule and teichoic acid to peptidoglycan, which are the last steps in connecting the different cell wall polymers to build the final cell wall. To obtain more insights into their cellular function, we first localized

TABLE 3. YWTF AND Cps2A CRYSTAL STRUCTURE PARAMETERS

	<i>Cps2A(R267A)–OprP</i>	<i>Cps2A(R267A)–OprPP</i>	<i>Ywtf</i>
Crystallization conditions	0.2 M diammonium citrate pH 5 15% PEG-3350	0.1 M sodium acetate pH 4.6 8% PEG-4K	0.1 M HEPES pH 7.5 20% PEG-3350 0.2 M MgCl ₂
Space group	P4 ₃ 2 ₁ 2	P4 ₃ 2 ₁ 2	P4 ₁ 2 ₁ 2
Unit cell (Å)	<i>a</i> = <i>b</i> = 73.19, <i>c</i> = 163.84	<i>a</i> = <i>b</i> = 73.50, <i>c</i> = 163.79	<i>a</i> = <i>b</i> = 65.74, <i>c</i> = 137.46
Resolution	19.72–1.95 (2.06–1.95)	40.95–1.77 (1.87–1.77)	47.5–1.79 (1.88–1.79)
Completeness	99.9 (99.9)	100 (100)	99.8 (100)
Multiplicity	15.7 (16.1)	13.8 (14.1)	6.1 (6.8)
Average I/signal	34.9 (8.8)	21.6 (7.0)	14.4 (2.9)
R _{sym}	0.054 (0.364)	0.08 (0.40)	0.069 (0.58)
R _{pim}	0.014 (0.092)	0.024 (0.109)	0.029 (0.24)
R _{cryst}	0.189	0.170	0.193
R _{free}	0.233	0.199	0.219
RMSD bond lengths (Å)	0.006	0.012	0.004
RMSD bond angles (°)	0.971	1.27	0.87
Average B (Å ²) protein	34.1	28.0	36.7
Average B (Å ²) lipid	41.9	49.4	64.0
Wilson B (Å ²)	27.4	20.5	29.1
Ramachandran plot	98.4/1.32	98.9/1.1	99.1/0.9
Favored/allowed (%)			
PDBid	4DE8	2XXQ	4DE9

fluorescent fusions of the three LCP proteins in *S. pneumoniae* using a recently published plasmid tool.¹³

Cellular localization of GFP-Cps2A, GFP-LytR, and GFP-Psr

To localize the three members of the LCP family in *S. pneumoniae*, the *cps2A*, *lytR*, or *psr* genes were cloned into pJWV25 in such a way that their 5'-ends were fused to *gfp*+.¹³ The resulting plasmids were transformed into the type 2 capsule strain D39 or the nonencapsulated strain R6. The expression of GFP+ -LCP fusion proteins was induced by the addition of Zn²⁺ to the growth medium. SDS-PAGE/western blot analyses of cell lysates confirmed that GFP+ fusion proteins of Cps2A, LytR, or Psr with the correct size were present in the lysate of the corresponding expression strains, detected with an anti-GFP antiserum, whereas no GFP band was visible in negative control lysates (Fig. 2A). Importantly, GFP bands with lower molecular weights were either not detectable, or had a low intensity, indicating that the expressed GFP-fusion proteins were not significantly degraded in the cell.

Next, we localized the GFP+ fusion proteins in D39 P_{czcD}-*gfp* + -*cps2A*, D39 P_{czcD}-*gfp* + -*psr*, and D39 P_{czcD}-*gfp* + -*lytR* strains by fluorescence microscopy.¹³ All three GFP-fusion proteins localized at the cell membrane (Fig. 2B), consistent with the predicted membrane anchor present in Cps2A, Psr, and LytR (Fig. 1B). Interestingly, all three proteins were enriched at the septal, mid-cell region of the cells, and this effect was most apparent for GFP+ -Psr (Fig. 2B). A similar localization pattern was observed in the nonencapsulated strain R6 (data not shown), indicating that the membrane localization of the three proteins, and their enrichment at the septum, does not depend on the presence of the capsule.

Capsule attachment in lcp mutant strains grown in liquid medium

In *B. subtilis*, the three LCP proteins have recently been implicated in the attachment of anionic polymers to the cell

wall peptidoglycan. Because the *B. subtilis* LCP proteins appear to have redundant roles *in vivo*, we aimed to test the contribution of each of the pneumococcal LCP proteins in the capsule attachment to the cell wall. For this, we have produced strains D39 *cps2A::ermB*, in which *cps2A* is replaced by an erythromycin cassette allowing the expression of the downstream capsule genes under their native promoter, and D39 *psr*, in which the *psr* gene is interrupted by the integrative plasmid pJDC9. Despite several attempts, we were unable to stably grow transformants of D39 *lytR::kan* in liquid C+Y growth medium (see next section). When growing in liquid medium, strain D39 *cps2A::ermB* contained a reduced amount (29%) of total capsule material compared with the parental D39 strain (Fig. 3), confirming previous studies.³¹ D39 *psr* contained slightly reduced amount of capsule (71%). Like the parental D39 strain, the *cps2A* and *psr* mutant strains did not release significant amounts of capsule into the supernatant, indicating that the capsule, although present in reduced amount, remains attached to the cell surface (Fig. 3). We next aimed to investigate the role of *lytR* in the attachment of capsule to the cell wall.

Capsule attachment in *cps2A lytR* cells

Inactivation of *lytR* in the nonencapsulated strain R6 leads to a severe phenotype where cells are mis-shapened and enlarged and often display misplaced septa.²² Transformation of the encapsulated strain D39 with DNA from R6 *lytR::kan* yielded small colonies but we were unable to propagate these cells in liquid culture. Therefore, we first aimed to use strains with a Zn²⁺-controlled *lytR* gene, allowing its repression in the absence of Zn²⁺ in the growth medium.¹³ We constructed D39 *cps2A::ermB lytR::kan* P_{czcD}-*lytR*, which contains the *lytR::kan* and *cps2A::ermB* mutations and, in addition, an ectopic copy of *lytR* at the *bgaA* site under the control of the Zn²⁺-inducible promoter P_{czcD}. However, this strain readily lost the repression of the ectopic *lytR* gene in the absence of Zn²⁺, resulting in constitutive expression of *lytR* from the *bgaA* site. Presumably, there is a

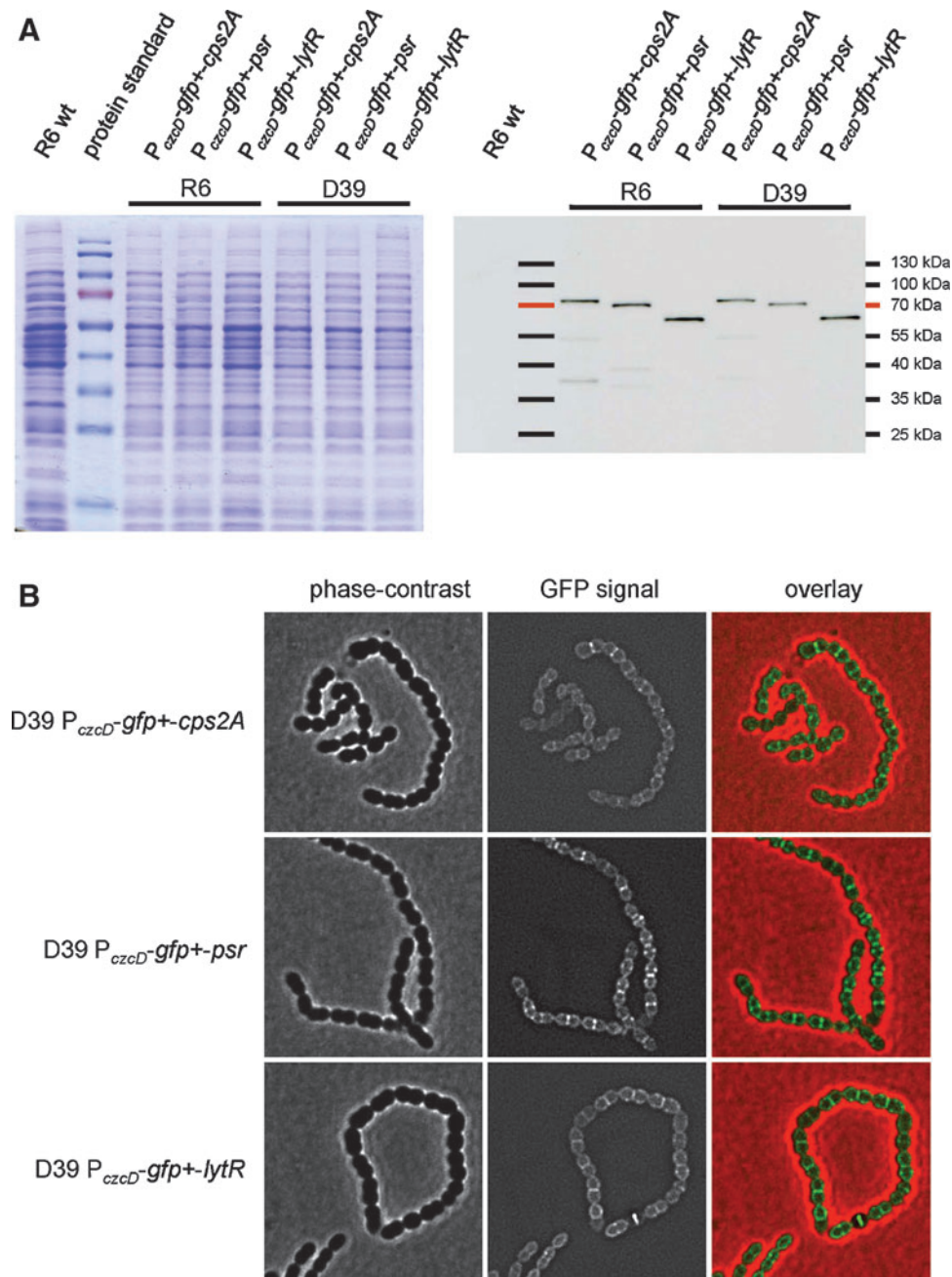


FIG. 2. Cellular localization of GFP+ -Cps2A, GFP+ -Psr, or Gfp+ -LytR. **(A)** Lysates of D39 or R6 cells expressing GFP+ fusion proteins were analyzed by SDS-PAGE (Coomassie staining, left side) and western blot using anti-GFP for detection of fusion proteins (right side). The GFP+ fusion proteins migrate according to their predicted molecular weight. **(B)** Fluorescence microscopy of D39 cells expressing GFP+ -Cps2A, GFP+ -Psr, or GFP+ -LytR. SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

strong selective pressure to generate suppressor mutations in the Zn^{2+} induction system leading to the expression of *lytR* without Zn^{2+} .

Because none of the encapsulated *lytR* mutants grew stably in liquid culture, we next quantified the amount of capsule in whole colonies obtained after transforming strains D39 or D39 *cps2A::ermB* with DNA from D39 *cps2A::ermB lytR::kan P_{czcD}-lytR* with selection for kanamycin resistance (Fig. 4). The state of the capsule in these transformants was also visualized by microscopy after staining with Congo

Red. Interestingly, while the D39 transformants were homogenous, the kanamycin-resistant D39 *cps2A::ermB* transformants produced two visibly different types of colonies that differed in the amount of capsule present (Fig. 4). The majority of transformants, herein referred to as D39 *cps2A::ermB lytR::kan*, were small and contained a low amount of, or perhaps no, capsule, comparable to that of R6 *lytR::kan* transformants. We were not able to grow cells from this colony type in liquid medium. At low frequency, we also obtained larger colonies containing capsule, although at a

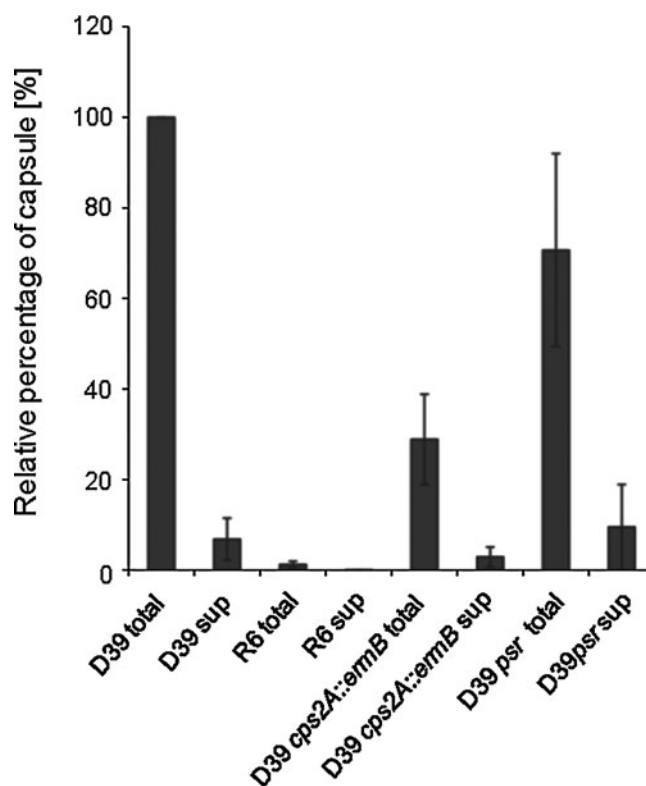


FIG. 3. The relative percentages of capsule in total cell lysates and in culture supernatants of various strains. Capsules were quantified by dot blot analysis using an anti-type 2 capsule antiserum. The values are the mean \pm SD of three independent samples.

reduced level to the D39 transformants. We conclude that these transformants have acquired a secondary, as yet unknown, mutation that allows the retention of the capsule in the *cps2A* *lytR* background. These transformants could be grown in liquid culture (see next section), and we refer to this suppressor strain as D39 *cps2A::ermB* *lytR::kan* (S). DNA from both types of kanamycin-resistant D39 *cps2A::ermB* transformants was extracted, and we confirmed by PCR analysis that the *cps2A* gene was inactivated by the erythromycin cassette. We also confirmed that the *bgaA* region was unchanged and the *lytR* gene could not be amplified. In addition, in both transformants, the *psr* gene sequence was identical to that in D39. Hence, both types of transformants were devoid of *cps2A* and *lytR* and had an intact *psr* gene, indicating that the differences in the amounts of capsules were not due to a difference in their LCP genes.

The cps2A lytR (S) suppressor mutant loses capsule material into the supernatant

The suppressor strain D39 *cps2A::ermB* *lytR::kan* (S) was able to grow in liquid C+Y growth medium, although growth was strongly impaired as indicated by the slower generation time compared with single *cps2A*, *lytR*, or *psr* mutant; the double *cps2A* *psr* mutant; and the parental strain (Fig. 5A). The *cps2A* *lytR* (S) suppressor mutant grew only to an optical density of 0.2–0.3 before lysing when inoculated from a fresh agar plate, and the strain did not grow at all when it was inoculated from liquid culture.

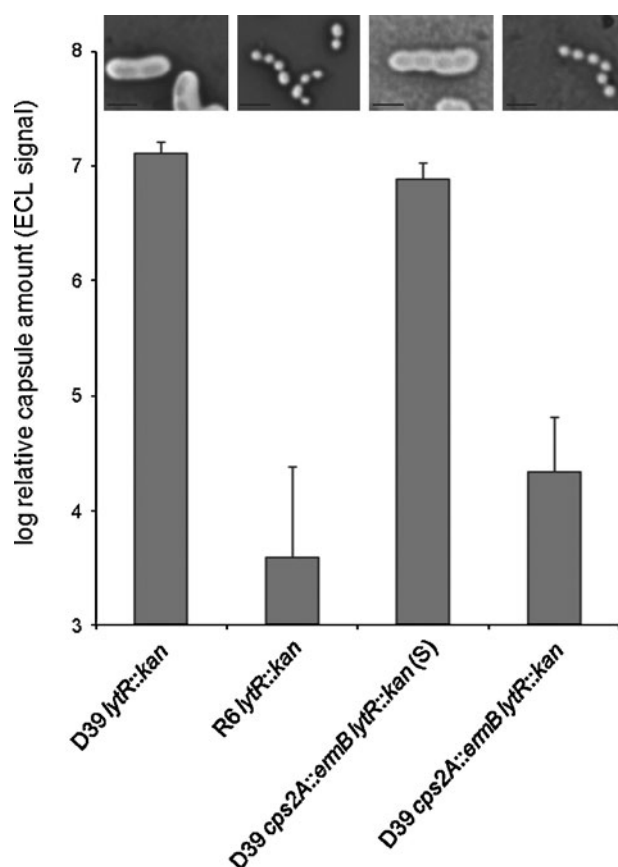


FIG. 4. Relative capsule amount in whole colonies of transformants. Strains were transformed to introduce the *lytR::kan* mutation. Whole colonies of transformants were taken and resuspended, and the cells were lysed prior to quantification of the capsule by dot blot analysis. The values are the mean \pm SD of three independent samples. The suppressor mutant D39 *cps2A::ermB* *lytR::kan* (S) contained far less capsule than D39 *lytR::kan*, whereas D39 *cps2A::ermB* *lytR::kan* and R6 *lytR::kan* have virtually no capsule. Micrograph insets above the histograms show capsule-stained cells of the strains.

We next quantified the capsule material in the supernatant of these cultures relative to the total capsule present in lysate (Fig. 5B). Interestingly, the D39 *cps2A::ermB* *lytR::kan* (S) suppressor strain released as much as 47% of its total capsule material into the supernatant, which was significantly more than the parental strain, the *cps2A* and *psr* single mutants, and the *cps2A* *psr* double mutant. The higher amount of capsule in the supernatant of D39 *cps2A::ermB* *lytR::kan* (S) was not due to lysis of cells, because the supernatant of this strain contained similar levels of proteins to those of the other strains (Fig. 5C).

When grown in liquid medium, D39 *cps2A::ermB* *lytR::kan* (S) produced 31% of the total capsule amount of D39 (data not shown), which is a similar reduction to that seen when the strain is grown on agar plates (Fig. 4). Indeed, electron micrographs of thin sections of cells from liquid culture show a strong reduction in surface-attached capsule compared with strain D39 (Fig. 6). Further, cells of D39 *cps2A::ermB* *lytR::kan* (S) were inhomogeneous in size and had misplaced septa, similar to R6 *lytR* mutant cells. Taken together, these

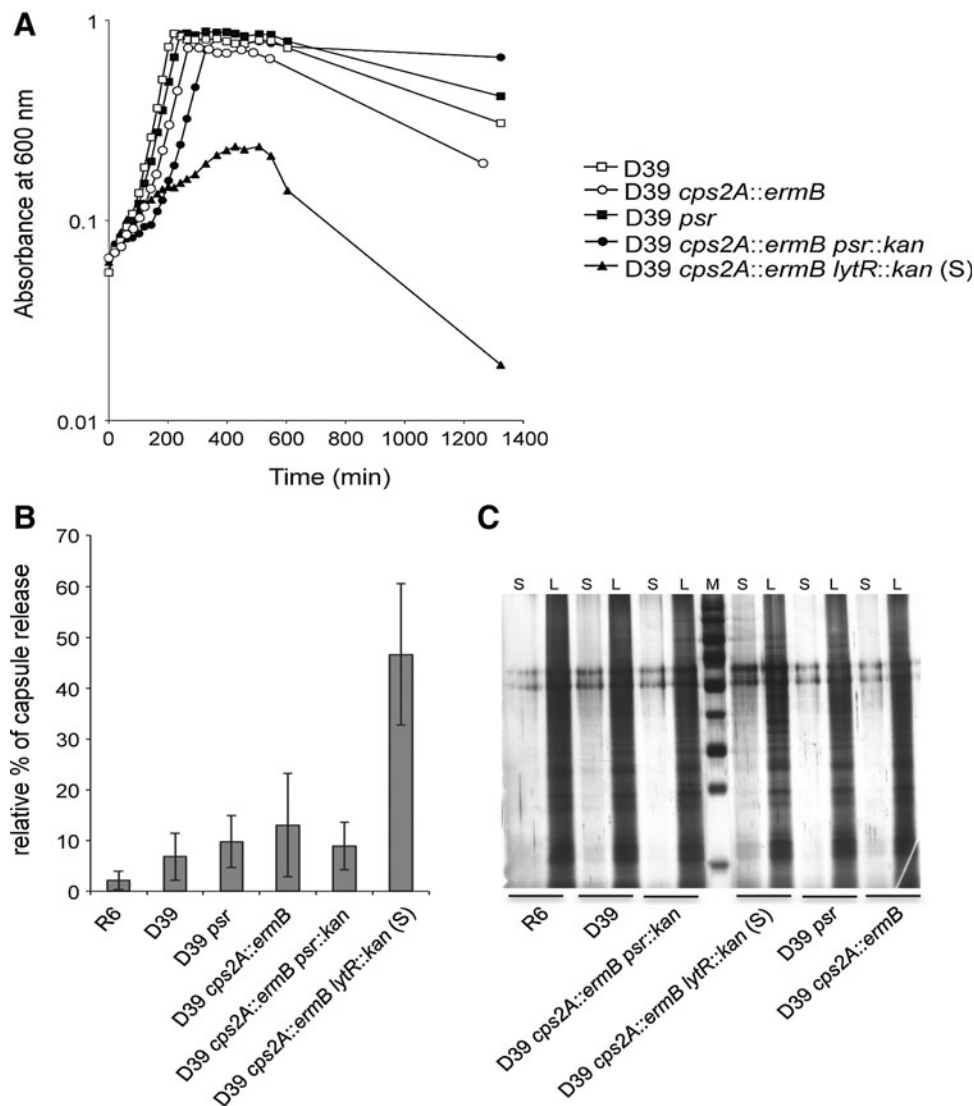


FIG. 5. Characterization of the suppressor mutant D39 *cps2A::ermB lytR::kan* (S). **(A)** Growth of strains in C + Y at 37°C. D39 *cps2A::ermB lytR::kan* (S) only reaches an OD₆₂₀ of 0.25 before autolysis begins. **(B)** Relative capsule amount in culture lysate and supernatant detected by dot blot analysis. D39 *cps2A::ermB lytR::kan* (S) loses ~47% of its capsule material into the culture supernatant. **(C)** Proteins detected in silver-stained SDS-PAGE of the lysate (L) and supernatant (S) samples analyzed in **(B)**.

data show that LytR is required for normal cell morphology and that it contributes with Cps2A to retaining the capsule at the cell surface.

Crystallographic evidence for polyprenol pyrophosphatase activity of Cps2A

In a previous report we showed that Cps2A and the *B. subtilis* homolog YwtF have a pyrophosphatase activity toward polyprenoid pyrophosphate lipid substrates, a reaction resembling the transfer of teichoic acid or capsular polysaccharides onto peptidoglycan.²⁴

We have now obtained crystallographic evidence that corroborates the ability of Cps2A to hydrolyze the phosphate-phosphate bond in polyprenol pyrophosphate lipids. When the extracellular domain of Cps2A is overexpressed in *E. coli*, it co-purifies with endogenous polyisoprenoid

phosphate lipids. Such Cps2A preparations have been crystallized as complexes with endogenous octaprenyl-pyrophosphate and decaprenyl-phosphate lipids.²⁴ We have subsequently found that the crystallization conditions affect the nature of the lipid bound to the protein. Identical batches of the Cps2A(R267A) mutant were crystallized under similar, but slightly different, conditions (0.1 M sodium acetate [pH 4.6] 8% PEG-4K vs. 0.2 M di-ammonium citrate [pH 5] 17% PEG-3350) and both crystal forms had the same space group with near-identical unit cell dimensions. In the structure of the wild-type Cps2A, Arg267 interacts with the lipid β -phosphate. The Cps2A(R267A) mutant protein proved to be, serendipitously, particularly amenable to crystallography. The major lipid in this preparation of Cps2A(R267A) was shown by mass spectroscopy to be octaprenyl-pyrophosphate.²⁴ However, in the refined crystal structures (refinement statistics in Table 3), the lipid composition of the

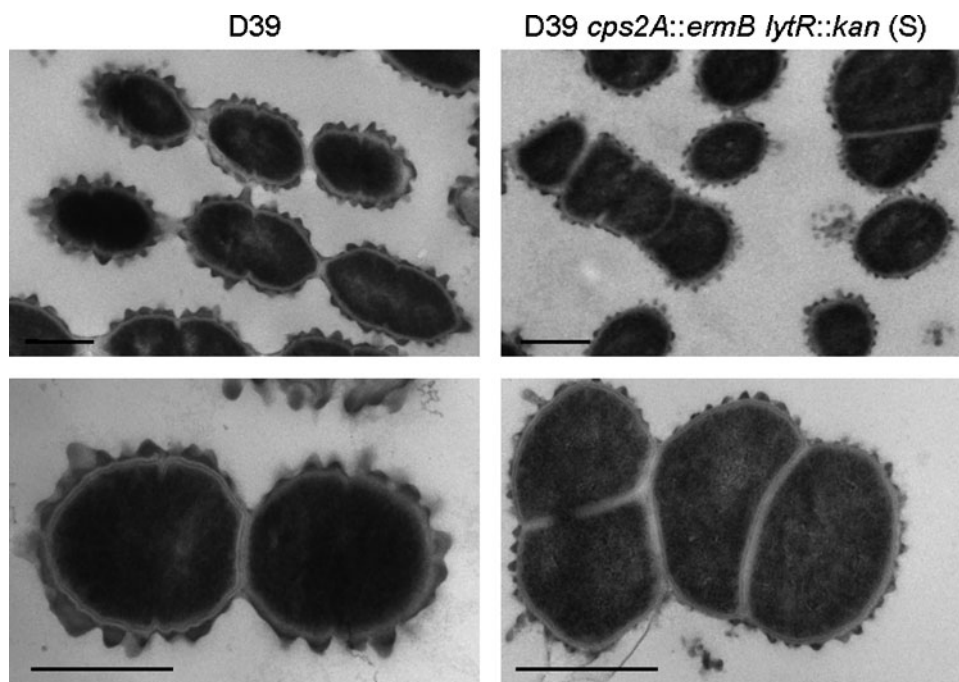


FIG. 6. Visualization of surface-attached capsule in D39 and D39 *cps2A::ermB lytR::kan* (S). Electron micrographs show reduced amount of cell-surface-associated capsule and aberrant septum placement and cell size. Bar, 0.5 μm .

crystallized protein differed. The crystals grown in acetate buffer clearly show a pyrophosphate lipid head group, even in electron density maps with phases calculated from only the protein atoms (Fig. 7A). In the crystals grown from citrate buffer, only a monophosphate lipid is visible. Further, an arginine residue, Arg242, has a significantly different position in the two crystal forms, which can be attributed to a salt-bridging interaction between Arg242 and the lipid β -phosphate.

As the crystals form over a much longer timescale in citrate buffer, not even appearing under a microscope for a month, the difference in lipid content is likely to be attributable to a change in the lipid content of the sample occurring during crystallization, presumably because of the pyrophosphatase activity described previously for Cps2A.²⁴ The Cps2A(R267A) mutant does indeed have pyrophosphatase activity in the di-ammonium citrate buffer used for crystallization. Using a colorimetric PiPer assay, a sample of 190 μM Cps2A(R267A) released $67 \pm 1.5 \mu\text{M}$ of inorganic phosphate after 5 days of incubation at 22°C in 20 mM di-ammonium citrate (pH 5.0). It might be argued that the Cps2A(R267A) sample contains a pre-existing mixture of lipids and the change in conditions favors crystallization of a particular protein-lipid complex from the mixture. However, it is hard to reconcile such a dramatic change in behavior with such a slight change in buffer pH and with the identical crystal packing that exists in the two crystal forms. Further, there are no obvious interactions of buffer ions with the lipid-binding site that could explain any buffer-specific stabilization of either protein-lipid complex. A change in lipid composition over the lengthy timescale of crystallization in citrate buffer, due to the innate pyrophosphatase activity of Cps2A, therefore, appears to be the most plausible explanation for this major change in crystallization behavior.

The B. subtilis LCP protein YwtF binds polyisoprenoid lipids

To simplify analysis of the proposed enzymatic function of LCP proteins, we sought to identify a homolog that could be isolated with an empty lipid-binding site. Structures of seven other LCP family proteins are currently in the PDB, none of which report a bound lipid molecule, though polyethylene glycol fragments have been built in the hydrophobic tunnel in PDB entry 3OWQ. The *B. subtilis* homolog of Cps2A, YwtF, was purified using a method similar to that of Cps2A and subjected to sparse matrix crystallization screening. Crystals were formed under identical conditions to those reported in the PDB (entry 3MEJ) and diffracted to a maximum resolution of 1.79 Å. 3MEJ was used as a starting model for refinement and was refined to an R_{free} of 0.24. This structure determination revealed a ribbon of electron density that corresponds to the bound lipid molecule when the structure of the Cps2A-lipid complex is superimposed on YwtF.

Electrospray mass spectrometry was subsequently used to assess whether the sample contained bound polyisoprenoid lipids (Fig. 7C). Spectra were recorded under both native and denaturing conditions and were consistent with the mass of the intact protein, minus initiating methionine (31,883 Da). Additional peaks at 32,607 Da and 32,731 Da were uniquely present in the native spectra. The differences in mass relative to the peak from the intact protein, 724 and 849 Da, are consistent with octaprenyl-pyrophosphate and undecaprenyl-phosphate, respectively, lipids which have been identified to co-purify with Cps2A.²⁴ Octaprenyl-pyrophosphate provided a better fit to the electron density map in the hydrophobic tunnel than the undecaprenyl-phosphate; the crystal packing also did not allow sufficient space to accommodate

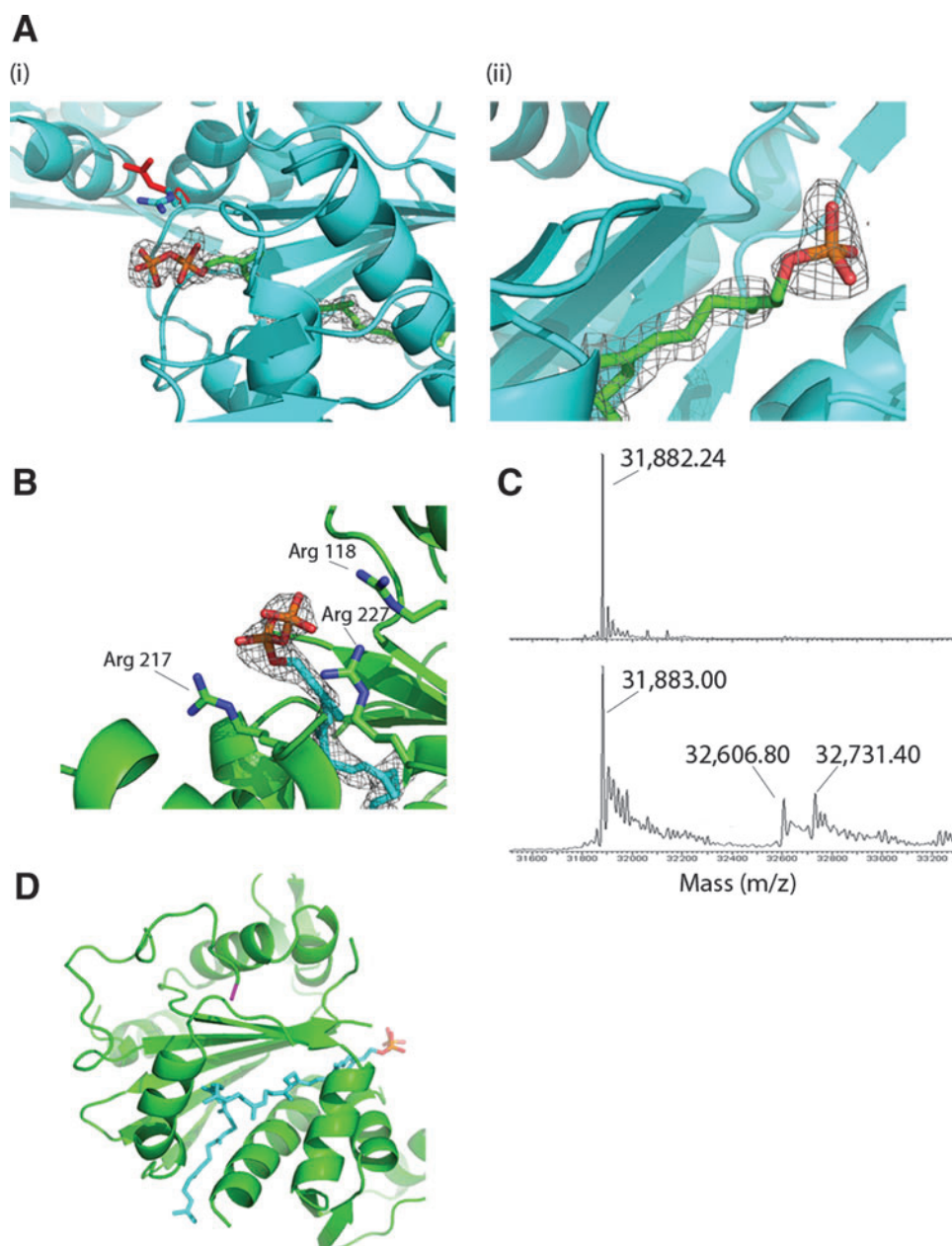


FIG. 7. Crystallographic evidence for pyrophosphatase activity of LCP proteins. **(A)** Effect of crystallization conditions on the lipid component of Cps2A(R267A) mutant. A complex with pyrophosphate lipid crystallizes in acetate buffer (i) and a complex with monophosphate lipid crystallizes in citrate buffer (ii). The position of Arg244 in the superimposed structures is shown in (i); the red color corresponds to Arg244 in the complex with monophosphate lipid. Both $2F_{\text{obs}} - F_{\text{calc}}$ maps are contoured at 1σ within 1.6 \AA of the bound lipid molecule; (i) shows an unbiased map calculated at an early stage of refinement, before the addition of lipid ($R_{\text{free}} = 0.27$); (ii) represents a map calculated with phases from a fully refined structure. **(B)** Structure of the complex between YwtF and octaprenyl-pyrophosphate. The conserved arginines 118, 217, and 242, which form salt bridges with the pyrophosphate head group, are shown. The $2F_{\text{obs}} - F_{\text{calc}}$ map is contoured at 1σ within 1.6 \AA of the bound lipid molecule. **(C)** Electrospray mass spectra of YwtF recorded under denaturing (upper) and native (lower) conditions. The spectra under native and denaturing conditions show a peak consistent with the mass of the expressed protein, minus the initiating methionine (31,883 Da). Additional peaks at 32,607 and 32,731 Da are unique to the spectra under native conditions; the mass differences relative to that of the protein-alone 31,883 Da (724 and 848 Da) are consistent with the masses of octaprenyl-pyrophosphate and undecaprenyl-pyrophosphate. **(D)** Model of undecaprenyl-pyrophosphate-YwtF complex. With the best fit of the lipid into the electron density and the steric constraints of the hydrophobic tunnel, the hydrophobic tail protrudes slightly into the solvent. However, the tail appears correctly orientated to interact with the membrane, given the position of the N-terminus of the protein (magenta), which bears the transmembrane helix.

the longer lipid. The lipid most convincingly filled the electron density in the all-*cis* configuration, rather than in the all-*trans* configuration (Fig. 7D), particularly for the isoprenoid group adjacent to the pyrophosphate. That the enzyme is stereospecific selective for lipid binding is underlined by fact that *trans*-polyprenoid lipids are more abundant in *E. coli* than *cis*-polyprenoid lipids, because of their roles as intermediates in quinone biosynthesis.⁴⁰ The electron density can accommodate di-*trans*, penta-*cis* octaprenyl-pyrophosphate, in which the two chiral double bonds furthest from the head group have the *trans* configuration. Di-*trans*, penta-*cis* octaprenyl-pyrophosphate is an intermediate in the synthesis of undecaprenyl-pyrophosphate⁴⁰ and may be sequestered preferentially by LCP proteins when expressed in the cytoplasm. Indeed, the final R_{free} in refinement was in actual fact at its lowest (0.22) when di-*trans*, penta-*cis* octaprenyl-pyrophosphate was modeled into the active site.

Attempts to model the longer lipid undecaprenyl-phosphate into the electron density resulted in a steric clash between the hydrophobic tail of the lipid and an adjacent protein molecule in the crystal lattice. In full-length YwtF the hydrophobic tail may extend out of the protein tunnel to interact with the membrane. This potential "exit" path from the hydrophobic lipid tunnel is reasonably orientated with respect to the N-terminus of the protein, to which the transmembrane helix is attached, for this to occur (Fig. 7D). Overall, these structural results support further an enzymatic role of LCP proteins in the transfer of phosphorylated chains of anionic cell wall polymers from an undecaprenyl-pyrophosphoryl-linked precursor to peptidoglycan on cleavage of the pyrophosphoryl bond.

Discussion

LCP proteins are widespread in Gram-positive bacteria and are often present in multiple versions in one species. *Lcp* mutant strains of various species show pleiotropic phenotypes and, therefore, LCP proteins have been suggested to be involved in the regulation of some aspects of cell wall metabolism. However, a recent publication²⁴ provided strong evidence that LCP proteins are not regulatory factors, but are in fact phosphotransferase enzymes for one of the last steps in cell wall assembly, the attachment of anionic cell wall polymers like capsular polysaccharides and teichoic acids, to peptidoglycan. This hypothesis is strongly supported by crystal structures of the extracytoplasmic part of pneumococcal Cps2A, which, depending on the *E. coli* expression strain from which it was produced, contained either a phosphorylated or pyrophosphorylated polyprenol lipid; these are the substrate and product mimics of the phosphotransferase reaction. The reaction presumably leads to the teichoic acid or capsule chains becoming attached to C6-OH of MurNAc residues in peptidoglycan by the formation of a phosphodiester bond. It is this last reaction in cell wall synthesis, connecting the major cell wall polymers with each other, which produces the final and functional cell wall in Gram-positive bacteria. Perhaps, it is therefore not surprising that affecting this reaction often leads to pleiotropic phenotypes as observed in *lcp* mutants.

The aim of this work was to investigate the role of the three pneumococcal LCP proteins in the attachment of the capsule to the cell wall. One of them, CpsA, is encoded by

the first gene of the large capsular gene cluster present in the different serotypes. A noticeable exception is serotype 3, which does not contain an *lcp* family gene in its capsule cluster and which does not covalently attach its capsule to peptidoglycan. Cps2A, the CpsA protein present in serotype 2 strains like D39, has been implicated previously in the attachment of capsule to the cell wall because a *cps2A* mutant strain has a reduced quantity (~30%–40%) of cell-wall-attached capsular polysaccharide. Because the residual capsule material is attached to the cell wall, we hypothesized that other LCP proteins, LytR and Psr, might contribute to capsule attachment in the absence of Cps2A. We also considered the possibility that Cps2A, LytR, and Psr attach teichoic acid chains to peptidoglycan to form wall teichoic acid, or attach teichoic acid chains to the glycolipid anchor to form lipoteichoic acid. Either there is one enzyme for each of these three attachment reactions, or LCP proteins have (semi)redundant roles in cell wall assembly. The latter possibility is supported by the available crystal structures of Cps2A and its *B. subtilis* homologs, which bind the lipid and pyrophosphoryl part of the anionic polymer precursor strongly. These structures also display a surface groove for binding peptidoglycan chains, but they do not appear to recognize much or any of the repeating chain of the anionic polymer. Hence, it is possible that some of these enzymes are "non-specific" in that they are able to attach different polyprenol pyrophosphoryl-linked polymers to peptidoglycan. This hypothesis is supported by the work presented in this article, which shows that LytR, at least, and perhaps Psr contribute to capsule-cell wall attachment in the absence of Cps2A. A role in cell wall assembly is also supported by the localization of GFP-fusions of all three LCP proteins at mid-cell (Fig. 2) where the pneumococcal cell wall grows. The fusion proteins also localized at mid-cell in the capsule-free strain R6 (not shown), indicating that capsule precursors are not required for their localization, supporting the hypothesis that LytR and Psr are active in cell wall assembly, presumably in the final assembly of teichoic acids. However, the precise role of each pneumococcal LCP protein in teichoic acid assembly remains to be determined.

D39 strains that lacked *cps2A* but expressed all other capsule genes had significantly reduced amounts of capsule material (Fig. 3), in confirmation of a previous report and indicating that LytR and Psr cannot complement fully the lack of Cps2A in the retention of the capsule. A *psr* mutant had slightly reduced amounts of capsule. Perhaps, Psr contributes to capsular assembly, or in this mutant Cps2A and LytR take over the Psr function to result indirectly in a reduced amount of capsule.

We were also unable to obtain a stable *lytR* mutant in the D39 background. *LytR* can be inactivated in the capsule-free R6 strain, leading to abnormal cell morphology and reduced growth. Presumably, the presence of a fully functional capsule assembly pathway adds to the growth problems of *lytR* mutants that prevent the isolation of viable mutants. Consistent with this premise, the inactivation of *lytR* in the D39 *cps2A::ermB* background, which contained less capsule material than D39 (Fig. 3), produced viable transformants. Interestingly, we have obtained two types of transformants, both of which were identical genetically, in that they both contained the *cps2A* and *lytR* deletions and retained an intact *psr*. The more abundant transformant type had little, if any,

capsule and was unable to grow in liquid medium; we infer that this is the true phenotype of a D39 *cps2A lytR* double mutant. The second type of transformant, less abundant and most likely a D39 *cps2A lytR* double mutant with one or more suppressor mutation(s), contained capsule and was able to grow in liquid medium, although at a slow growth rate and only for one culture cycle. Interestingly, these cells lost massive amounts of capsule material into the growth medium without showing significant lysis, indicating that the capsule material was secreted into the growth medium.

The most likely explanation for our results is as follows. All three LCP homologs have semiredundant roles in the attachment of capsular polysaccharide and teichoic acid precursor chains to peptidoglycan, and perhaps of teichoic acid precursor chains to the glycolipid anchor. Deleting one or two of the corresponding genes results in an imbalance in the assembly of cell wall polymers causing growth defects, in particular when *lytR* is deleted. This indicates that *LytR* has specific function(s) that cannot readily be taken over by the two homologs. It could be expected that *lcp* mutants accumulate lipid-linked teichoic acid or capsule precursors in their cell membrane. However, in several *lcp* mutant strains investigated, the lipid-linked precursors do not accumulate but the final cell wall polymer is present at a reduced level. We hypothesize that the cytoplasmic synthesis and polymerization of teichoic acid and capsular polysaccharide chains, their transport across the membrane, and their attachment to the cell wall peptidoglycan are coupled, explaining the reduced amount of anionic polymers in *lcp* mutant strains (Figs. 2, 5, and 6). In the D39 *cps2A lytR* suppressor strain, this feedback regulation appears to be lost and thus the strain produces an excess of capsule material that cannot be retained in the cell wall and is lost into the culture supernatant. We are currently investigating how the capsule material is released from the cell membrane in this strain, and which genes are mutated to produce this phenotype.

At the molecular level, the new structures reported here bolster the evidence that LCP proteins are the phosphotransferases that catalyze the transfer of anionic cell wall polymers (teichoic acids and capsular polysaccharides) from lipid-linked intermediates onto peptidoglycan. First, the head group of the lipid changes when the timescale of crystallization is extended, consistent with pyrophosphorolysis of the lipid occurring within a crystallization drop. This pyrophosphorolysis is consistent with our previous observation that inorganic phosphate accumulates in purified preparations of the Cps2A-octaprenyl-pyrophosphate complex.²⁴ The pyrophosphorolysis reaction has chemistry resembling that of the phosphotransfer of teichoic acid or capsule onto peptidoglycan.

Second, we have also shown that the *B. subtilis* LCP protein YwtF binds lipids that resemble the substrate of the phosphotransfer reaction. The presence of exogenous *E. coli* octaprenyl-pyrophosphate lipid in the crystallized protein, in addition to undecaprenyl-phosphate, was confirmed by mass spectroscopy. On examination of the structure of YwtF bound to octaprenyl-pyrophosphate, many of the interactions between charged residues and the lipid head group resemble those seen in Cps2A.²⁴ The only exception is Asp82 (equivalent to in Asp234 in Cps2A). In Cps2A, Asp234 coordinates a magnesium ion that interacts with the pyrophosphate head group, whereas in YwtF, Asp82 is more

distant from the phosphate and is at the start of a disordered loop region that could not be modeled. In Cps2A, this loop region is in close proximity to an additional protein domain that is unique to Cps2A. The proximity of this domain may stabilize the loop by restraining its conformational mobility, which may in turn enable Asp234 to coordinate the magnesium ion. The disorder in this loop in YwtF may explain why there is no magnesium bound to the pyrophosphate head group in the YwtF structure, despite the presence of 0.2 M MgCl₂ in the crystallization conditions. Modulation of the mobility of this loop due to interactions with other components of the cell wall synthetic machinery could in fact provide a mechanism to regulate the activity of the enzyme.

The disorder in this loop may also explain the apparent weaker binding of the lipid to YwtF. In comparing the crystallographic B factors of the lipid to those of the surrounding protein, the octaprenyl-pyrophosphate appears to be more mobile in the YwtF structure (Table 3). Further, purified preparations of YwtF catalyze the pyrophosphorolysis of exogenously added geranyl pyrophosphate but Cps2A, purified in the same manner, is inactive in the same assay.²⁴ These observations suggest that the endogenous *E. coli* lipids are more loosely bound to YwtF than in Cps2A, and that the active site is, therefore, more accessible to exogenously added substrates.

In summary, our genetic data have provided insights into the complex mechanisms of capsule attachment in *S. pneumoniae*, pointing to a semiredundant role of all three LCP proteins in this process. Our structural data have reinforced the proposed role of LCP proteins as the enzymes that catalyze the transfer of teichoic acid and/or capsular polysaccharides onto peptidoglycan. Although the precise role each of these enzymes plays in cell wall assembly has yet to be defined, their inactivation would seem likely to impair pneumococcal growth and/or survival at an infection site. Therefore, we consider Cps2A and its homologs as valid targets in the search for antimicrobials that target pneumococci and, potentially, other Gram-positive pathogens.

Acknowledgments

This work was supported by a grant from the UK Biotechnology and Biological Sciences Research Council to W.V. and R.J.L. The authors thank the Electron Microscopy Research Service of Newcastle University for sample preparation, the Mass Spectroscopy Service of Leeds University's Astbury Centre for mass spectroscopy, and Dr. Arnaud Balsé of Newcastle University and the staff of beamlines at Diamond Light Source for help during diffraction data collection.

References

1. Adams, P.D., P.V. Afonine, G. Bunkóczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.-W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, and P.H. Zwart. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst.* **D66**:213–221.
2. Avery, O.T., C.M. Macleod, and M. McCarthy. 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of

- transformation by desoxyribonucleic acid fraction isolated from pneumococcus type III. *J. Exp. Med.* **79**:137–157.
3. Baddiley, J. 1972. Teichoic acids in cell walls and membranes of bacteria. *Essays Biochem.* **8**:35–77.
 4. Bender, M.H., and J. Yother. 2001. CpsB is a modulator of capsule-associated tyrosine kinase activity in *Streptococcus pneumoniae*. *J. Biol. Chem.* **276**:47966–47974.
 5. Bentley, S.D., D.M. Aanensen, A. Mavroidi, D. Saunders, E. Rabbino-witsch, M. Collins, K. Donohoe, D. Harris, L. Murphy, M.A. Quail, G. Samuel, I.C. Skovsted, M.S. Kalltoft, B. Barrell, P.R. Reeves, J. Parkhill, and B.G. Spratt. 2006. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. *PLoS Genet.* **2**:e31.
 6. Bogaert, D., R. De Groot, and P.W. Hermans. 2004. *Streptococcus pneumoniae* colonisation: the key to pneumococcal disease. *Lancet Infect. Dis.* **4**:144–154.
 7. Bratcher, P.E., K.H. Kim, J.H. Kang, J.Y. Hong, and M.H. Nahm. 2010. Identification of natural pneumococcal isolates expressing serotype 6D by genetic, biochemical and serological characterization. *Microbiology* **156**:555–560.
 8. Bui, N.K., A. Eberhardt, D. Vollmer, T. Kern, C. Bougault, A. Tomasz, J.P. Simorre, and M. Vollmer. 2012. Isolation and analysis of cell wall components from *Streptococcus pneumoniae*. *Anal. Biochem.* **421**:657–666.
 9. Calix, J.J., and M.H. Nahm. 2010. A new pneumococcal serotype, 11E, has a variably inactivated *wcjE* gene. *J. Infect. Dis.* **202**:29–38.
 10. Chen, J.D., and D.A. Morrison. 1987. Cloning of *Streptococcus pneumoniae* DNA fragments in *Escherichia coli* requires vectors protected by strong transcriptional terminators. *Gene* **55**:179–87.
 11. Crisostomo, M.I., W. Vollmer, A.S. Kharat, S. Inhülsen, F. Gehre, S. Buckenmaier, and A. Tomasz. 2006. Attenuation of penicillin resistance in a peptidoglycan O-acetyl transferase mutant of *Streptococcus pneumoniae*. *Mol. Microbiol.* **61**:1497–1509.
 12. Denapaite, D., R. Brückner, R. Hakenbeck, and W. Vollmer. 2012. Biosynthesis of teichoic acid in *Streptococcus pneumoniae* and closely related species: lessons from genomes. *Microb. Drug. Resist.* **18**:344–358.
 13. Eberhardt, A., L.J. Wu, J. Errington, W. Vollmer, and J.W. Veening. 2009. Cellular localization of choline-utilization proteins in *Streptococcus pneumoniae* using novel fluorescent reporter systems. *Mol. Microbiol.* **74**:395–408.
 14. Emsley, P., and K. Cowtan. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**:2126–2132.
 15. Evans, P. 2006. Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62**:72–82.
 16. Frolet, C., M. Beniazza, L. Roux, B. Gallet, M. Noirclerc-Savoye, T. Vernet, and A.M. Di Guilmi. 2010. New adhesin functions of surface-exposed pneumococcal proteins. *BMC Microbiol.* **10**:190.
 17. Guidolin, A., J.K. Morona, R. Morona, D. Hansman, and J.C. Paton. 1994. Nucleotide sequence analysis of genes essential for capsular polysaccharide biosynthesis in *Streptococcus pneumoniae* type 19F. *Infect. Immun.* **62**:5384–5396.
 18. Hakenbeck, R., A. Madhour, D. Denapaite, and R. Brückner. 2009. Versatility of choline metabolism and choline-binding proteins in *Streptococcus pneumoniae* and commensal streptococci. *FEMS Microbiol. Rev.* **33**:572–586.
 19. Hanson, B.R., B.A. Lowe, and M.N. Neely. 2011. Membrane topology and DNA-binding ability of the Streptococcal CpsA protein. *J. Bacteriol.* **193**:411–420.
 20. Henrichsen, J. 1995. Six newly recognized types of *Streptococcus pneumoniae*. *J. Clin. Microbiol.* **33**:2759–2762.
 21. Henriques, M.X., T. Rodrigues, M. Carido, L. Ferreira, and S.R. Filipe. 2011. Synthesis of capsular polysaccharide at the division septum of *Streptococcus pneumoniae* is dependent on a bacterial tyrosine kinase. *Mol. Microbiol.* **82**:515–534.
 22. Johnsborg, O., and L.S. Havarstein. 2009. Pneumococcal LytR, a protein from the LytR-CpsA-Psr family, is essential for normal septum formation in *Streptococcus pneumoniae*. *J. Bacteriol.* **191**:5859–5864.
 23. Kabsch, W. 2010. Xds. *Acta Crystallogr. D Biol. Crystallogr.* **66**:125–132.
 24. Kawai, Y., J. Marles-Wright, R.M. Cleverley, R. Emmins, S. Ishikawa, M. Kuwano, N. Heinz, N.K. Bui, C.N. Hoyland, N. Ogasawara, R.J. Lewis, W. Vollmer, R.A. Daniel, and J. Errington. 2011. A widespread family of bacterial cell wall assembly proteins. *Embo J.* **30**:4931–4941.
 25. Kern, T., M. Giffard, S. Hediger, A. Amoroso, C. Giustini, N.K. Bui, B. Joris, C. Bougault, W. Vollmer, and J.P. Simorre. 2010. Dynamics characterization of fully hydrated bacterial cell walls by solid-state NMR: evidence for cooperative binding of metal ions. *J. Am. Chem. Soc.* **132**:10911–10919.
 26. Lanie, J.A., W.L. Ng, K.M. Kazmierczak, T.M. Andrzejewski, T.M. Davidsen, K.J. Wayne, H. Tettelin, J.I. Glass, and M.E. Winkler. 2007. Genome sequence of Avery's virulent serotype 2 strain D39 of *Streptococcus pneumoniae* and comparison with that of unencapsulated laboratory strain R6. *J. Bacteriol.* **189**:38–51.
 27. Lazarevic, V., P. Margot, B. Soldo, and D. Karamata. 1992. Sequencing and analysis of the *Bacillus subtilis* *lytRABC* divergon: a regulatory unit encompassing the structural genes of the N-acetylmuramoyl-L-alanine amidase and its modifier. *J. Gen. Microbiol.* **138**:1949–1461.
 28. Leslie, A.G. 2006. The integration of macromolecular diffraction data. *Acta Crystallogr. D Biol. Crystallogr.* **62**:48–57.
 29. McCoy, A.J., R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, and R.J. Read. 2007. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**: 658–674.
 30. Morona, J.K., J.C. Paton, D.C. Miller, and R. Morona. 2000. Tyrosine phosphorylation of CpsD negatively regulates capsular polysaccharide biosynthesis in *Streptococcus pneumoniae*. *Mol. Microbiol.* **35**:1431–1442.
 31. Morona, J.K., R. Morona, and J.C. Paton. 2006. Attachment of capsular polysaccharide to the cell wall of *Streptococcus pneumoniae* type 2 is required for invasive disease. *Proc. Natl. Acad. Sci. U S A* **103**:8505–8510.
 32. Morrison, D.A. 1997. Streptococcal competence for genetic transformation: regulation by peptide pheromones. *Microb. Drug Resist.* **3**:27–37.
 33. Morrison, D.A., S.A. Lacks, W.R. Guild, and J.M. Hageman. 1983. Isolation and characterization of three new classes of transformation-deficient mutants of *Streptococcus pneumoniae* that are defective in DNA transport and genetic recombination. *J. Bacteriol.* **156**:281–290.
 34. Murshudov, G.N., A.A. Vagin, and E.J. Dodson. 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**:240–255.
 35. Nelson, A.L., A.M. Roche, J.M. Gould, K. Chim, A.J. Ratner, and J.N. Weiser. 2007. Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. *Infect. Immun.* **75**:83–90.
 36. Ottolenghi, E., and R.D. Hotchkiss. 1960. Appearance of genetic transformic activity in pneumococcal cultures. *Science* **132**:1257–1258.

37. Park, I.H., D.G. Pritchard, R. Cartee, A. Brandao, M.C. Brandileone, and M.H. Nahm. 2007. Discovery of a new capsular serotype (6C) within serogroup 6 of *Streptococcus pneumoniae*. J. Clin. Microbiol. 45:1225–1233.
38. Sorensen, U.B., J. Henrichsen, H.C. Chen, and S.C. Szu. 1990. Covalent linkage between the capsular polysaccharide and the cell wall peptidoglycan of *Streptococcus pneumoniae* revealed by immunochemical methods. Microb. Pathog. 8:325–334.
39. Sung, C.K., H. Li, J.P. Claverys, and D.A. Morrison. 2001. An *rpsL* cassette, janus, for gene replacement through negative selection in *Streptococcus pneumoniae*. Appl. Environ. Microbiol. 67:5190–5196.
40. Teng, K.H., and P.H. Liang. 2011. Structures, mechanisms and inhibitors of undecaprenyl diphosphate synthase: a cis-prenyltransferase for bacterial peptidoglycan biosynthesis. Bioorg. Chem. [Epub ahead of print]; DOI:10.1016/j.bioorg.2011.09.004
41. Tomasz, A. 1964. Studies on the competence (for genetic transformation) of *Diplococcus pneumoniae* in a synthetic medium. Abstract presented at the 64th Annual Meeting of the American Society for Microbiology. American Society for Microbiology. Washington D.C. Abstract no. G87.
42. Vagin, A., and A. Teplyakov. 2010. Molecular replacement with MOLREP. Acta Crystallogr. D Biol. Crystallogr. 66:22–25.
43. van der Poll, T., and S.M. Opal. 2009. Pathogenesis, treatment, and prevention of pneumococcal pneumonia. Lancet 374:1543–1556.
44. Vollmer, W. 2008. Structural variation in the glycan strands of bacterial peptidoglycan. FEMS Microbiol. Rev. 32:287–306.
45. Vollmer, W., and A. Tomasz. 2000. The *pgdA* gene encodes for a peptidoglycan N-acetylglucosamine deacetylase in *Streptococcus pneumoniae*. J. Biol. Chem. 275:20496–20501.
46. Vollmer, W., and A. Tomasz. 2002. Peptidoglycan N-acetylglucosamine deacetylase, a putative virulence factor in *Streptococcus pneumoniae*. Infect. Immun. 70:7176–7178.
47. Vollmer, W., D. Blanot, and M.A. de Pedro. 2008. Peptidoglycan structure and architecture. FEMS Microbiol. Rev. 32:149–167.
48. Weidenmaier, C., and A. Peschel. 2008. Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. Nat. Rev. Microbiol. 6:276–287.

Address correspondence to:

Dr. Waldemar Vollmer

Centre for Bacterial Cell Biology

Institute for Cell and Molecular Biosciences

Newcastle University

Richardson Road

Newcastle upon Tyne NE2 4AX

United Kingdom

E-mail: w.vollmer@ncl.ac.uk

Richard J. Lewis, M.A., D.Phil.

Centre for Bacterial Cell Biology

Institute for Cell and Molecular Biosciences

Newcastle University

Richardson Road

Newcastle upon Tyne NE2 4AX

United Kingdom

E-mail: rick.lewis@ncl.ac.uk

A widespread family of bacterial cell wall assembly proteins

Yoshikazu Kawai¹, Jon Marles-Wright¹,
Robert M Cleverley¹, Robyn Emmins¹,
Shu Ishikawa², Masayoshi Kuwano²,
Nadja Heinz¹, Nhat Khai Bui¹, Christopher
N Hoyland¹, Naotake Ogasawara²,
Richard J Lewis¹, Waldemar Vollmer¹,
Richard A Daniel¹ and Jeff Errington^{1,*}

¹Centre for Bacterial Cell Biology, Medical School, Newcastle University, Newcastle upon Tyne, UK and ²Graduate School of Information Science, Nara Institute of Science and Technology, Nara, Japan

Teichoic acids and acidic capsular polysaccharides are major anionic cell wall polymers (APs) in many bacteria, with various critical cell functions, including maintenance of cell shape and structural integrity, charge and cation homeostasis, and multiple aspects of pathogenesis. We have identified the widespread LytR–Cps2A–Psr (LCP) protein family, of previously unknown function, as novel enzymes required for AP synthesis. Structural and biochemical analysis of several LCP proteins suggest that they carry out the final step of transferring APs from their lipid-linked precursor to cell wall peptidoglycan (PG). In *Bacillus subtilis*, LCP proteins are found in association with the MreB cytoskeleton, suggesting that MreB proteins coordinate the insertion of the major polymers, PG and AP, into the cell wall.

The EMBO Journal (2011) 30, 4931–4941. doi:10.1038/emboj.2011.358; Published online 30 September 2011

Subject Categories: cell & tissue architecture; microbiology & pathogens

Keywords: capsular polysaccharide; LytR–CpsA–Psr proteins; MreB; peptidoglycan; teichoic acids

Introduction

The cell wall is crucial for maintaining the structural integrity and the characteristic shape of a bacterial cell. In Gram-positive bacteria, a group that contains many important pathogens, the cell wall has two major components: (i) peptidoglycan (PG), whose pathway is the target for some of our most successful antibacterial compounds (the β -lactams and glycopeptides) and (ii) the PG-attached anionic cell wall polymers (APs), including wall teichoic acids (WTAs) and acidic polysaccharides.

WTAs and their lipid-linked versions (lipoteichoic acids) have a wide range of important cellular roles including

control of autolytic activity, antigenicity and innate immune recognition, pathogenicity, biofilm formation, efficient release of secreted proteins into the culture medium, cation homeostasis, antibiotic resistance, and cell elongation and division (summarized in Weidenmaier and Peschel, 2008).

In *Bacillus subtilis*, WTA is present in large quantities, roughly equal to those of PG (Foster and Popham, 2002). Most steps in the synthesis of WTA are now known, catalysed by a variety of gene products with a *tag* prefix (Neuhaus and Baddiley, 2003; Brown *et al*, 2008). The polymer is synthesized in the cytoplasm, translocated across the membrane by the ABC transporter TagGH (Schirner *et al*, 2011), and is covalently attached to PG outside the cell (Yokoyama *et al*, 1989). The physical connection of major APs like WTA or capsular polysaccharides to PG builds the final cell wall architecture and is essential for proper cell wall functionality, but the enzyme catalysing this important step remains to be identified.

Disruption of most of the *tag* genes of *B. subtilis* is lethal. This lethality can be suppressed by deleting *tagO* or *tagA* to block the earliest steps in the pathway, which presumably prevents either the build up of a toxic intermediate or the sequestration of an essential metabolite. Nevertheless, *tagO* and *tagA* mutants are severely compromised in cell growth and they lose the ability to elongate, becoming rounded and swollen (D'Elia *et al*, 2006a,b, 2009). *tagO* null mutants of the round *Staphylococcus aureus* have no significant growth defect but they are affected in spatial control of PG synthesis, are hypersensitive to lysozyme, and are badly impaired in virulence (Weidenmaier and Peschel, 2008; Atilano *et al*, 2010).

In most rod-shaped bacteria, cell elongation is governed largely by the prokaryotic actin homologue, MreB, which assembles into patches or filaments at the inner surface of the membrane (Graumann, 2009). MreB is essential for viability in many bacteria, and its depletion induces severe morphological defects (Jones *et al*, 2001; Figge *et al*, 2004; Kruse *et al*, 2005; Slovak *et al*, 2005; Hu *et al*, 2007). Many Gram-positive bacteria have more than one MreB isoform. The *B. subtilis* MreB isoforms, MreB, Mbl, and MreBH largely colocalize and have partially overlapping functions (Carballido-Lopez *et al*, 2006; Kawai *et al*, 2009a). MreB and Mbl are essential under normal growth conditions, although the mutants are viable in high Mg^{2+} concentrations for reasons that are not yet clear (Formstone and Errington, 2005; Schirner and Errington, 2009). Evidence from several laboratories supports the idea that the MreB cytoskeleton somehow spatially regulates the synthesis of PG and potentially other cell wall components, thereby bringing about controlled expansion of the wall, while retaining a defined cell shape (Daniel and Errington, 2003; Vats *et al*, 2009; Kawai *et al*, 2009b).

We have now obtained several lines of evidence supporting the idea of a direct role for the MreB cytoskeleton in WTA synthesis and cell wall attachment. By searching for MreB-interacting proteins, we have identified a widely distri-

*Corresponding author. Centre for Bacterial Cell Biology, Medical School, Newcastle University, Richardson Road, Newcastle upon Tyne NE2 4AX, UK. Tel.: +44 191 222 8126; Fax: +44 191 222 7424; E-mail: jeff.errington@ncl.ac.uk

Received: 9 August 2011; accepted: 31 August 2011; published online: 30 September 2011

buted family of proteins of previously unknown function, the LytR–Cps2A–Psr (LCP) family, as having a role in the biogenesis of WTA and anionic polysaccharides (Hübscher *et al*, 2008). LCP proteins are present in virtually all gram-positive bacteria, which characteristically contain PG-attached WTA and/or APs. Our structural and biochemical data on several LCP proteins provide strong evidence that this family of proteins carry out the key step of attaching APs to the cell wall PG. The LCP proteins present an important novel antibiotic target and may aid our understanding of assembly of heteropolymeric cell walls in fungi and plants.

Results

A role for MreB proteins in WTA synthesis or assembly

B. subtilis mutants affected in genes for lateral wall PG synthesis (*mreC*, *mreD*, *rodA*, and a double mutant of *pbpA* and *pbpH*) have a common phenotype in which the cells become spherical (Henriques *et al*, 1998; Wei *et al*, 2003; Leaver and Errington, 2005), and their growth requires high concentrations of Mg^{2+} (Supplementary Figure S1A). Mutants lacking all three *mreB* genes also adopt a spherical shape (Schirner and Errington, 2009; Kawai *et al*, 2009a) but they are not rescued by Mg^{2+} (Supplementary Figure S1A). Mutants affected in WTA synthesis are also spherical (Lazarevic and Karamata, 1995; Soldo *et al*, 2002; Bhavsar *et al*, 2004). We noticed that *tagG* and *tagH* mutants were, like the *mreB* triple mutant, unresponsive to Mg^{2+} (Supplementary Figure S1B). We hypothesized that the *mreB* triple mutant might be Mg^{2+} -unresponsive because the MreB cytoskeleton does not only control the Mg^{2+} -responsive PG synthesis, but additionally the Mg^{2+} -unresponsive WTA synthesis. To test this hypothesis, we made use of the unique feature of mutants affected in the later steps of WTA synthesis, which can be rescued by disruption of the *tagO* gene (D'Elia *et al*, 2006a,b). If this was correct, mutation of *tagO* might also suppress the lethality of triple *mreB* mutation. Figure 1A shows that disruption of *tagO* suppresses lethality in a conditional mutant in which the one remaining *mreB* homologue is under IPTG control, supporting the idea that the MreB proteins are required for one or more late steps in WTA synthesis.

The MreB cytoskeleton associates with known and novel components of the WTA biosynthetic machinery

We used *in vivo* cross-linking, followed by IMAC affinity copurification of oligohistidine-tagged MreB, Mbl, or MreBH, and liquid chromatography tandem mass spectrometry analysis (LC-MS/MS) to identify proteins associated with the MreB cytoskeleton. The list of proteins that pulled down with the cytoskeleton (summarized in Figure 2A, with full data for each of the isoforms in Supplementary Table SI) included various proteins already known to associate with MreB, such as PBP1, PBP2A, PBP4, PBPH, and MreC (Kawai *et al*, 2009a,b). Importantly, the list also included several proteins of the WTA pathway: TagD, TagE, TagF, and TagH, indicative of a role for the cytoskeleton in the synthesis or the assembly of WTA.

The WTA synthetic pathway of *B. subtilis* is the best characterized of any organism. However, although WTA polymer is known to be covalently linked to PG (Yokoyama *et al*, 1989), the transferase required for this reaction has not

been identified. If the MreB protein was responsible for spatial organization of WTA, this factor should also be present in the MreB complexes. We analysed the list of candidates focusing on three key parameters: transmembrane (TM) topology, with a major (likely catalytic) domain predicted to be located outside the cytoplasm; gene location in proximity to known WTA synthetic loci (because WTA systems are variable and can be exchanged *en bloc* by genetic transformation) (Young *et al*, 1989); and phylogenetic distribution largely matching that of WTA. This analysis highlighted three members of the poorly characterized LCP family of genes and proteins (Figure 2). All three *B. subtilis* members of this gene family (*lytR*, *ywtF*, and *yvhJ*) are located within a 50-kbp segment of DNA that contains virtually all of the known WTA genes (Figure 2B). The three genes were renamed *tagT* (*ywtF*), *tagU* (*lytR*), and *tagV* (*yvhJ*).

Bacterial two-hybrid analysis revealed strong interactions between MreB and the TagT and TagU proteins (Supplementary Figure S2A). GFP fusions were made to each protein and expressed from the *amyE* locus under the control of a xylose-inducible promoter. GFP-TagU at least was functional based on a complementation tests (see below). Localization patterns of GFP-TagT and -TagU were reminiscent of those of MreB proteins, with patterns of distributed dots and tilted bands. These patterns were disrupted by double mutation of *mreB* and *mbl*, and by dissolution of the membrane potential, which Strahl and Hamoen (2010) recently showed was required for localization of MreB and associated proteins (Supplementary Figure S2C). GFP-TagV gave a less clear-cut pattern and the localization was not affected by *mreB* mutation or treatment with CCCP (Supplementary Figure S2C). These observations again support the idea that TagT and TagU, at least, are associated with the MreB cytoskeleton.

Triple disruption of the tagTUV genes causes loss of rod shape and is lethal under normal growth conditions

Single mutants of *tagT*, *tagU*, or *tagV* had no discernible effect on cell growth or morphology under normal growth conditions. To test for possible functional redundancy, all combinations of double mutants were then generated and confirmed to be viable. $\Delta tagTU$ and $\Delta tagUV$ mutants did not differ noticeably from the wild type (Figure 1B–D). However, the $\Delta tagTV$ mutant grew more slowly than the wild type and the cells were generally wider and often showed abnormal bulging (Figure 1E). A triple disruption of the *tagTUV* genes was apparently not viable. To investigate the phenotype of a triple mutant, *tagV* was placed under the control of the IPTG-inducible P_{spac} promoter in a $\Delta tagTU$ double mutant background. The resultant strain grew as normal rod-shaped cells in the presence of IPTG (Figure 1I). However, when IPTG was removed, growth was arrested and the morphology of the cells changed, becoming shorter and wider with noticeable abnormal bulging (Figure 1G), eventually becoming almost round (Figure 1H). Thus, at least one functional TagT, TagU, or TagV protein is required for viability and normal cell morphogenesis.

Disruption of the tagTUV genes can be suppressed by tagO mutation but not Mg^{2+}

If the *tagTUV* triple mutant was defective in WTA synthesis or assembly, it should be rescued by deletion of *tagO* but not by high Mg^{2+} concentration. As anticipated, growth of the triple

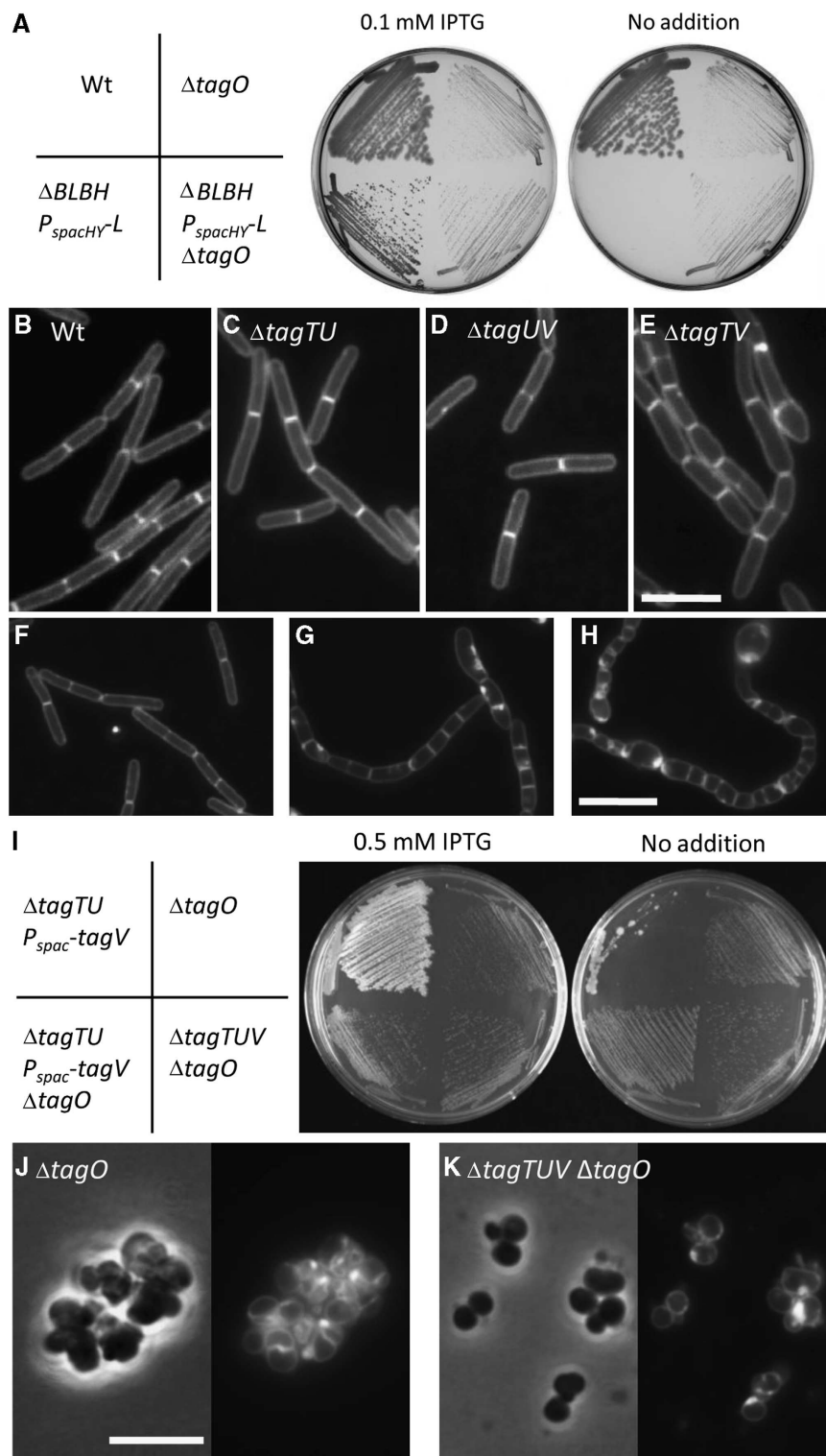


Figure 1 Lethal phenotype of a triple mutant of *mreB* paralogues and *tagTUV*, and rescue by disruption of *tagO*. (A) Growth of strains 168 (wild-type), YK992 ($\Delta tagO$), YK1190 (*amyE::P_{spacHY}-mbl ΔmreB Δmbl ΔmreBH ΔtagO*), and YK1119 (*amyE::P_{spacHY}-mbl ΔmreB Δmbl ΔmreBH*) on NA plates with or without 0.1 mM IPTG. (B–E) Cell morphologies of typical fields of strains 168 (wild-type, B), RE204 ($\Delta tagTU$, C), YK915 ($\Delta tagUV$, D), and YK917 ($\Delta tagTV$, E). The cell membranes were stained with Nile Red. Scale bar represents 5 μm. (F–H) Cell morphologies of typical fields of strain YK914 ($\Delta tagTU P_{spac-tagV}$) cultured in the presence (F) or absence (G, H) of 0.5 mM IPTG. Images were taken at 180 (G) and 240 (H) min after removal of IPTG. The cell membranes were stained with Nile Red. Scale bar represents 5 μm. (I) Growth of strains YK1031 ($\Delta tagTU P_{spac-tagV}$ pMAP65), YK1030 ($\Delta tagTU P_{spac-tagV} \Delta tagO$), YK1033 ($\Delta tagTUV \Delta tagO$), and YK1163 ($\Delta tagO$) on LB plates with or without 0.5 mM IPTG. (J, K) Cell morphologies of typical fields of strains YK1163 ($\Delta tagO$, J) and YK1033 ($\Delta tagTUV \Delta tagO$, K). The cell membranes were stained with Nile Red (right). Scale bar represents 5 μm.

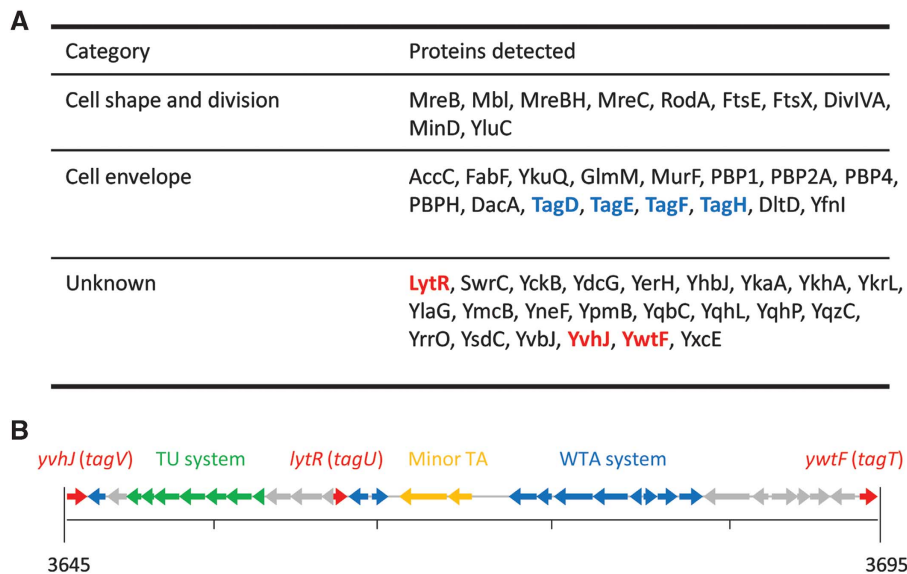


Figure 2 The WTA biosynthetic machinery and LCP family of proteins associated with the MreB cytoskeleton. **(A)** Summary of proteins associated with MreB cytoskeleton (see also full data in Supplementary Table S1). WTA synthetic and LCP proteins are indicated in blue and red, respectively. **(B)** Genetic organization of three AP systems in *B. subtilis*, the WTA biosynthetic genes (blue), the teichuronic acid (TU) biosynthetic genes (green), and the minor TA biosynthetic genes (yellow). *lytR* homologues, *tagT*, *U*, and *V*, were indicated in red. Numbers show the position on the *B. subtilis* chromosome.

tagTUV mutant was not rescued by culture in the presence of various concentrations of Mg^{2+} up to 25 mM (Supplementary Figure S1C). We then took a null mutation of *tagO* and combined this with disruptions of the *tagT* *U* and *V* genes and *P_{spac}-tagV*. The resulting strain grew even in the absence of IPTG (Figure 1I), suggesting that disruption of *tagO* suppresses the lethality of *tagTUV* triple disruption. To confirm this, we were able to construct the quadruple mutant directly, taking advantage of the proximity of the *tagV* and *tagO* genes (see Supplementary data). The resulting strain was viable, and its morphological phenotype was indistinguishable from that of a *tagO* single mutant (Figure 1J and K).

TagTUV proteins are required for WTA synthesis or assembly

To test directly for a WTA defect, cell wall material (the insoluble PG-WTA complex) was isolated from various strains and acid treated to solubilize the WTA polymers, which were subsequently separated by PAGE and visualized by alcian blue-silver staining (Wolters *et al*, 1990) (Figure 3A). Extracts of the wild-type strain revealed the expected ladder-like pattern, representing partially hydrolysed WTA polymers with a range of distinct chain lengths (lanes 1 and 6). WTA material was completely absent from cells of the *tagO* mutant, in which the first step in the WTA pathway is blocked (lane 5). The staining pattern for the single *tagU* mutant was indistinguishable from that of the wild type (lane 2). When the triple *tagTUV* mutant containing the P_{spac} -*tagV* construct was grown in the presence or absence of IPTG, the amount of WTA polymer was greatly reduced in the absence of IPTG (lane 4) compared with its presence (lane 3), consistent with *tagV* being required for WTA accumulation in the wall. Residual WTA present in lane 4 is probably due to incomplete repression of the P_{spac} promoter in the samples.

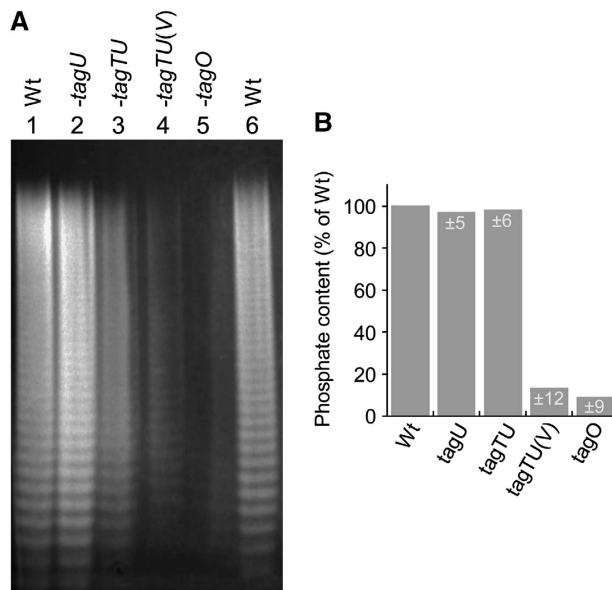


Figure 3 Effects of *tagTUV* mutants on WTA synthesis or assembly. **(A, B)** Cells of 168 (wild-type, lanes 1 and 6), RE201 ($\Delta tagU$, lane 2), YK914 ($\Delta tagTUV P_{spac-tagV}$, lanes 3 and 4), and YK992 ($\Delta tagO$, lane 5) were cultured with (lane 3) or without (lanes 1, 2 and 4–6) IPTG. Purified WTA samples were separated and visualized as described in the Materials and methods **(A)**. Phosphate content of cell wall was assayed as described in the Materials and methods **(B)**. White numbers indicate standard deviation from three independent experiments.

Cell wall phosphate, which is present only in WTA and is not found in PG, was also quantified in the various mutants (Figure 3B). Consistent with a previous report (Soldo *et al*, 2002; D'Elia *et al*, 2009), the wall phosphate content of the *tagO* null mutant was greatly reduced compared with wild-type cells. By contrast, the wall phosphate content of the *tagU* single mutant or the *tagTUV* triple mutant containing the

P_{spac}-tagV construct grown in the presence of IPTG had almost wild-type cell wall phosphate contents. However, in the absence of IPTG, wall phosphate was greatly reduced, almost to the levels found in the *tagO* mutant. We therefore concluded that TagTUV proteins have a critical function in the formation of a WTA-loaded cell wall.

The structure of Δ TM-Cps2A from *Streptococcus pneumoniae*

Attempts to develop direct biochemical assays for TagTUV activity, using mimics of the complex lipid-linked teichoic acid substrate, were not successful in the first instance. To obtain molecular insights into the LCP family of proteins, and their likely enzymatic function, we screened a number of LCP proteins lacking the N-terminal, TM spanning region (Δ TM-) for crystallization. We solved the structure of the Δ TM version of the capsule synthesis protein CpsA from serotype 2 *S. pneumoniae* D39 (named Δ TM-Cps2A from here on) by selenomethionine SAD and refined the subsequent atomic model at 1.69 Å resolution (Supplementary Table SII).

Δ TM-Cps2A comprises two distinct domains: domain 1 (the accessory domain) comprises residues 111–213 and domain 2 (the LCP domain) spans residues 214–481 (Figure 4). The structure of the accessory domain, the sequence (and indeed the presence) of which is not conserved in the wider LCP family of proteins, is depicted in

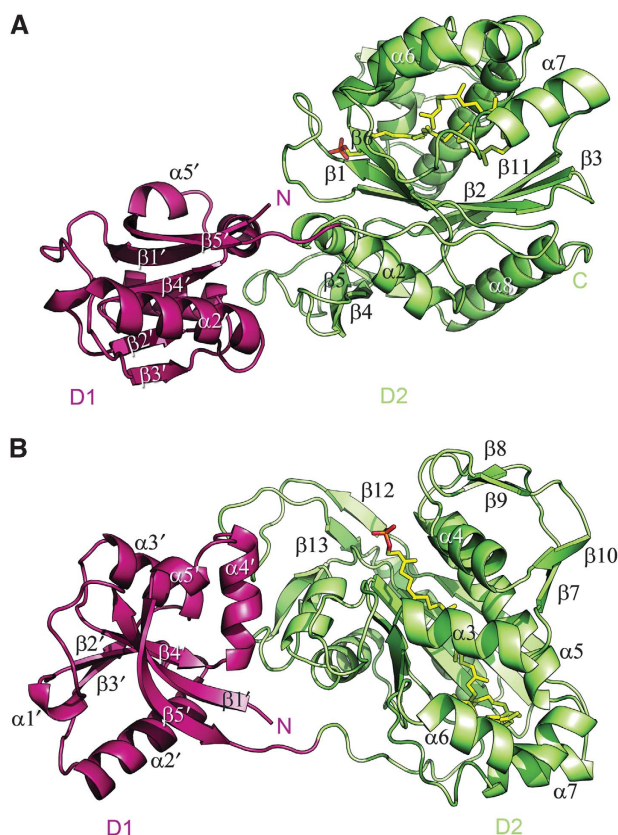


Figure 4 The structure of Δ TM-Cps2A. (A, B) Orthogonal views of the extracellular portion of Cps2A, shown as a cartoon, with the accessory domain coloured red and the LCP domain coloured green. The decaprenyl-phosphate present in the active site is shown as a stick model, with carbon atoms coloured yellow, phosphorous in orange, and oxygen in red. Secondary structure elements in both domains are labelled from N- to C-termini of each domain.

Figure 4A and B, and described in detail in Supplementary Figure S3. The LCP domain of Δ TM-Cps2A, which shares $\sim 30\%$ sequence identities with the equivalent domains from the three *B. subtilis* LCP proteins, is described below.

The LCP domain binds phosphorylated polyisoprenoid lipids

The LCP domain has an α - β - α architecture with a five-stranded β -sheet forming the core of the protein and α -helices surrounding the sheet on both faces (Figure 4). Two pairs of β -strands ($\beta 4/\beta 5$ and $\beta 12/\beta 13$) extend away from the protein core to form the interface between the two domains of Δ TM-Cps2A, burying some 570 Å² of surface area in the process. Crucially, a polyisoprenoid phosphate lipid was found in a hydrophobic pocket between the main β -sheet and α -helices $\alpha 3$ – $\alpha 7$ (Figure 4). Presumably, Δ TM-Cps2A had bound the lipid when heterologously expressed in the *Escherichia coli* host, consistent with its affinity for a lipid-linked capsule precursor in *S. pneumoniae*. The lipid was built as monotrans, octa-*cis* decaprenyl-phosphate (dpr-*P*).

We were subsequently able to solve the structure of Δ TM-Cps2A by molecular replacement in the presence of all *cis* octaprenyl-pyrophosphate (opr-PP) bound in the lipid-binding site. Again, no effort was made to load the protein with lipid prior to crystallization. The variation in lipid content between preparations of Δ TM-Cps2A probably represents a difference in the lipid composition of the two *E. coli* strains used: BL21 (DE3) leads to opr-PP being bound whereas B834 (DE3) results in dpr-*P* being bound. In both cases, the identity of the bound lipid was determined by electrospray mass spectroscopy of protein:lipid complexes. Mass differences between free proteins and lipid-bound proteins were 774 and 724 Da (Supplementary Figure S4A and B), corresponding to dpr-*P* and opr-PP (equivalent molecular masses 777 and 723 Da, respectively). The electron density maps prior to the building of the bound lipids, and during refinement, are consistent with the presence of dpr-*P* in one structure and opr-PP in the other. Both lipids are found naturally in bacteria (Bouhss *et al*, 2008).

The polyisoprenoid-binding pocket is lined with hydrophobic side chains (Figures 4 and 5A and C) from residues that, although not strictly identical, are completely conserved in hydrophobic character across the entire LCP family of proteins. The cavity for the lipid increases in diameter beyond the sixth isoprenoid, so much so that the lipid is folded back upon itself where the cavity for lipid binding is at its widest, and there are fewer interactions with the protein (Figure 5A and C). Consequently, the electron density in the region of the sixth prenoic moiety in both the dpr-*P* and opr-PP complexes is relatively poor, as is the electron density for the terminal prenoic group in the dpr-*P* structure. The key interactions between lipid and protein are, however, maintained in both structures (Figure 5A and C).

The phosphate headgroup of the dpr-*P* lipid is held in place by a number of conserved, charged residues. The invariant R267, R362, and R374 form key interactions with all the phosphate oxygens to form a positively charged pocket in the protein surface (Figure 5B). The conserved D371 and Q378 are in contact with these arginine residues to stabilize further their conformations.

The lipid-proximal phosphate in the opr-PP-bound structure is shifted slightly and now makes contacts to D234,

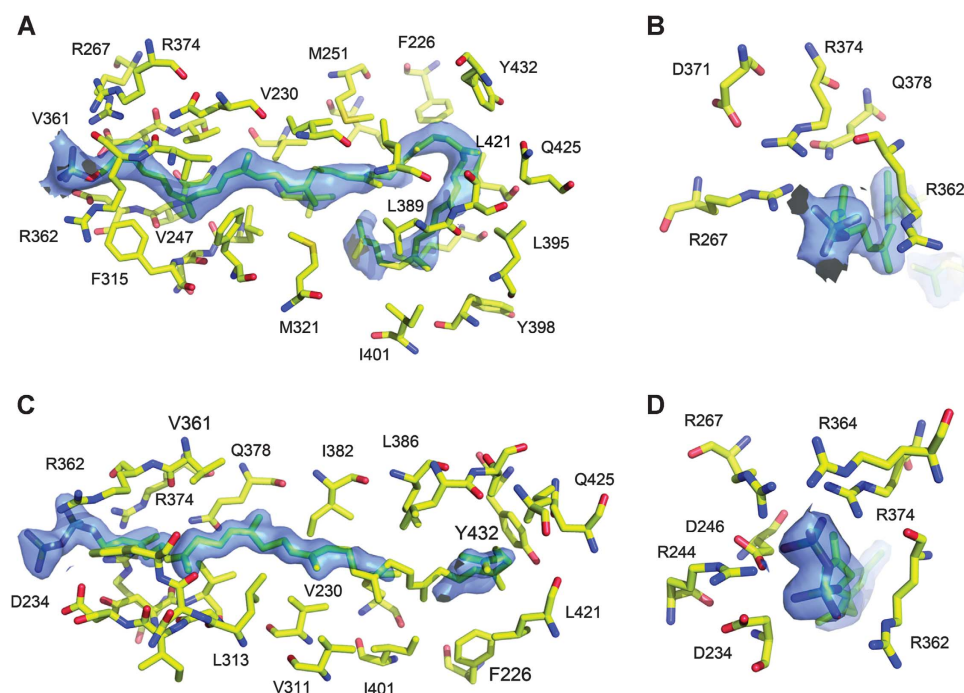


Figure 5 Lipid binding by the LCP domain of Δ TM-Cps2A. (A, B) decaprenyl-phosphate binding to Δ TM-Cps2A and (C, D) octaprenyl-pyrophosphate bound Δ TM-Cps2A. Protein residues within 4 Å of the lipid are shown as stick models, with final, σ^A -type $2mF_{obs}-DF_{calc}$ electron density maps shown as blue transparent surfaces and contoured in both cases at 1.0σ .

R244, and R362, whereas the lipid-distal phosphate contacts the invariant arginines, R244, R267, R364, and R374, fixing the position and orientation of the pyrophosphate (Figure 5D). In the dpr-*P* structure, R267 is built in two conformations, only one of which interacts with the phosphate. Similarly, in the dpr-*P* structure, R244 and R364 point into solvent, but, as with R267, readjust their conformations in the presence of a pyrophosphate lipid headgroup.

Since the structure solution of Δ TM-Cps2A, seven unpublished outputs from the Northeast Structural Genomics Consortium have been deposited in the PDB that likely represent structures of other LCP proteins. None of these entries have been characterized biochemically or genetically. These structures (PDBids: 3PE5 (from *Clostridium leptum*, UNIPROT code A7VV38); 3QFI (*Enterococcus faecalis*, Q83812); 3OWQ (*Listeria innocua*, Q92CZ6); 3NXH (*B. subtilis*, P96499); 3NRO (*Listeria monocytogenes*, Q8Y889); 3OKZ (*Streptococcus algalactiae*, Q8E703), and 3MEJ (*B. subtilis*, Q7WY78)) are the only meaningful homologues to Δ TM-Cps2A that can be found in the PDB by either secondary structure matching or amino-acid sequence searches and they are mostly, and apparently incorrectly, annotated as transcriptional regulators, whereas there is no significant structural homology of these structures, or of Δ TM-Cps2A, to known transcription factors in the PDB. The range of sequence identities of the seven hits, in comparison to the catalytic domain of Cps2A, is between 29 and 32% with BLAST *E*-values between 5^{-23} and 8^{-16} . All but one (3PE5) can be superimposed on Δ TM-Cps2A with r.m.s.d.s of 1.8–2.2 Å on 200–230 matched C α atoms of the catalytic domain (for 3PE5, the r.m.s.d. is 2.8 Å on 224 matched C α s). None of these seven PDBs contain bound lipid, despite the close sequence and structural homology in global terms, and in the regions of the proteins that bind the lipid. The

hydrophobic residues in Δ TM-Cps2A that are within van der Waals' contact of the bound isoprenoid lipids have hydrophobic amino acids in the equivalent positions in the other available structures. The charged amino acids in Δ TM-Cps2A in the vicinity of the pyrophosphate headgroup are almost completely invariant in the other structures (Supplementary Figure S5). Consequently, it is not evident why Δ TM-Cps2A should purify from *E. coli* with lipids bound, whereas the other LCP proteins do not.

The LCP proteins likely catalyse a missing step in cell wall synthesis

Anionic (or acidic) polymers, for instance teichoic acids or capsular polysaccharides, are synthesized inside the cell, and are attached to a carrier lipid, undecaprenyl phosphate for translocation ('flipping') to the outside of the cell. In the final stage of the biosynthesis of the cell wall of bacteria, teichoic acids are covalently attached to PG, whereby the structure of the acceptor (PG precursor, nascent PG chain, or cross-linked PG) is not known. This phosphotransfer reaction links the phospho-teichoic acid chain to PG MurNAc residues releasing undecaprenyl phosphate PG whereby. Our structural and genetic data suggest that the LCP proteins catalyse this step.

The observation of polyisoprenyl (pyro)phosphate lipids bound to Δ TM-Cps2A identified the active site for the entire LCP family (Supplementary Table SIII). The roles of key, conserved amino acids in the lipid-binding pocket were confirmed by mutagenesis of the *B. subtilis* tagU gene (see below and Figure 6A).

LCP proteins are magnesium-dependent phosphotransferases

The importance of the two aspartic acids was clarified by our discovery that D234 and D246 coordinate a magnesium ion

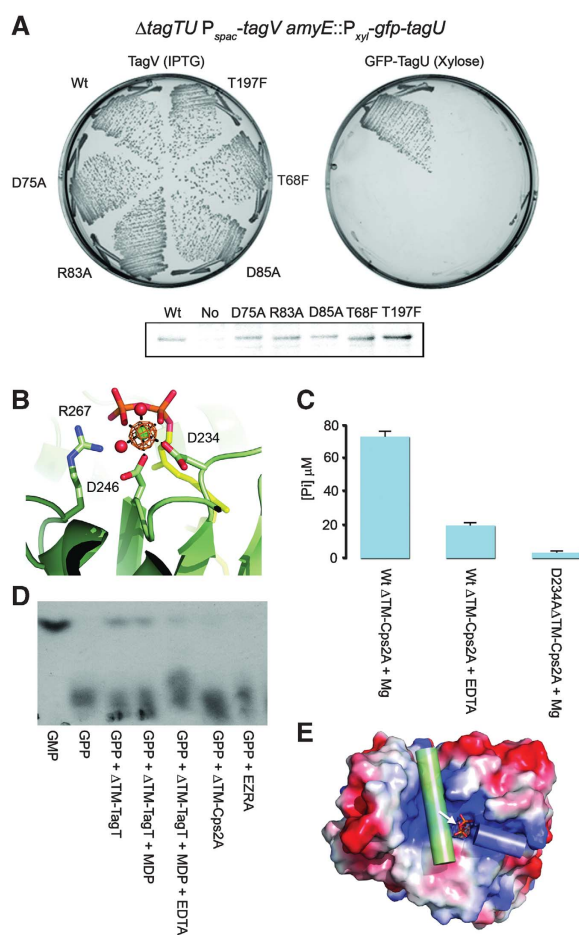


Figure 6 The biochemical properties of LCP proteins. (A) Growth of strains YK1254 (wild-type GFP-TagU), YK1365 (D75A), YK1366 (R83A), YK1367 (D85A), YK1368 (T86F), and YK1369 (T197F) on NA plates with 0.5 mM IPTG or 0.5% xylose. Western blot analysis of GFP-TagU levels in strain expressing wild-type and various mutant GFP-TagU fusions with or without (No) 0.5% xylose indicate correct accumulation of the proteins and consequently mutations D75A, R83A, D85A, T86F, and T197F in TagU are unable to sustain growth in the absence of wild-type copy of *tagT*, *tagU*, and *tagV*. (B) Metal binding by ΔTM -Cps2A. Crystals of Cps2A soaked in manganese show a clear peak in an anomalous difference map calculated for data collected at the manganese *K*-edge (orange mesh). Metal coordinating residues are shown as ball and stick representations, with the position of R267A shown with blue carbons. (C) Incubation of ΔTM -CPS2A–octaprenyl-pyrophosphate complex with magnesium releases inorganic phosphate. Addition of EDTA, or mutation to alanine of aspartate 234, which coordinates the catalytic magnesium ion, both suppress release of inorganic phosphate. Bar length is proportional to the wild-type experiment with Mg^{2+} present ($76.31 \pm 2.9 \mu$ M released with Mg^{2+} present versus $20.60 \pm 1.26 \mu$ M released with EDTA present; 3.38 ± 0.68 released by D234A). (D) Pyrophosphatase activity of ΔTM -TagT. A purified ΔTM -TagT protein catalyses hydrolysis of the pyrophosphate bond in exogenously added geranyl pyrophosphate (ger-PP) to generate geranyl monophosphate (ger-P), monitored by thin layer chromatography. Parallel incubations with the unrelated protein, EzrA (purified using the same procedures as the LCP proteins herein) and ΔTM -Cps2A reveal no pyrophosphatase activity towards exogenous ger-PP; in the latter case, this can be attributed to the blockage of the active site by endogenous opr-PP lipid. Hydrolysis is magnesium dependent, but is not affected by the addition of the PG fragment, *N*-acetylmuramyl dipeptide (MDP). (E) Electrostatic protein surface of the LCP domain of ΔTM -Cps2A, with bound octaprenyl-pyrophosphate ligand. The scissile bond in the phosphotransferase reaction catalysed by this enzyme class is shown with an arrow. The potential pathway for AP binding is highlighted in blue, semi-transparent rod (labelled AP), and the likely position of PG, to which the phosphorylated AP will be transferred, is shown with the green rod (labelled PG).

that is situated between the two phosphate groups of opr-PP in a further structure of ΔTM -Cps2A (Figure 6B). The presence of opr-PP in the sample was confirmed by mass spectrometry. The assignment of magnesium in the structure was confirmed by a separate structure determination of ΔTM -Cps2A–R267A–opr-PP with manganese bound, where the position of the bound metal ion was verified in an anomalous difference Fourier (Figure 6B), and the identity of the bound lipid by mass spectrometry; Supplementary Figure S4C). The bound Mg^{2+} obeys near perfect octahedral geometry, with one ligand each from D234, D246 and oxygens from both phosphates (Figure 6B). The other two coordination positions are filled by water molecules. The side chain of R244 is displaced by the binding of divalent cations, and no longer interacts with the pyrophosphate headgroup. Otherwise, the interactions with the opr-PP lipid are the same as in the absence of the metal, including the abutment of the distal phosphate by the side chain of R267 (Figure 5D), which presumably aids catalysis by stabilizing the transition state.

The role of magnesium in catalysis was investigated with a colourimetric assay to measure the liberation of free phosphate from ΔTM -Cps2A that had been purified with opr-PP bound, to demonstrate the capacity of ΔTM -Cps2A to cleave the phosphoester bond in pyrophosphate-containing lipids. In this assay, the absence of Mg^{2+} ions from the reaction, or the presence of EDTA, produced significantly less P_i in an end point assay than when Mg^{2+} ions were present (Figure 6C). Similarly, removal of one of the key Mg^{2+} coordination partners, D234, by its mutation to alanine produced significantly less free P_i when incubated in the presence of Mg^{2+} (Figure 6C). The key role for the Mg^{2+} coordination by D234 is consistent with our observation that a *B. subtilis* TagU carrying the equivalent mutation, D75A, is not functional *in vivo* (Figure 6A; Supplementary Table SIII).

In contrast to ΔTM -Cps2A, purified preparations of *B. subtilis* ΔTM -TagT were not loaded with an *E. coli* lipid, as confirmed by X-ray crystallography, and catalysed the hydrolysis of the pyrophosphate phosphoester bond in an exogenously added lipid, geranyl pyrophosphate (ger-PP), in a magnesium-dependent manner (Figure 6D). The addition of high molecular weight PG, or of a small PG fragment, *N*-acetylmuramyl dipeptide, did not stimulate the transfer of the terminal phosphoryl group to the possible acceptor. Most likely, the reaction catalysed by the LCP proteins favours phosphotransfer from AP-loaded substrates (which are not commercially available), over that from the unloaded diphosphoryl lipid substrate present in our reaction. Nonetheless, our *in vivo* and *in vitro* data strongly suggest that the LCP family of proteins is the missing enzyme that catalyse the linkage of teichoic acids and other acidic/APs to PG.

Discussion

LCP proteins are phosphotransferases

Studying the final step of bacterial cell wall synthesis, the attachment of TA to PG, has been proven very challenging to multiple laboratories for several reasons: (i) the substrates have highly complex, heterogeneous structures that are not commercially available, (ii) the precise identity of the PG acceptor (precursor, nascent chain, or cross-linked polymer) is not known, and (iii) the reaction is catalysed at the

cytoplasmic membrane by integral membrane proteins and is likely linked to membrane transport of the precursors. Although we could not reconstitute this complex system *in vitro*, in this paper, we present data that, in combination, strongly suggest that members of the LCP family of proteins are the long searched for enzymes that attach APs to cell wall PG. This conclusion is based on the following observations: (i) the TagTUV paralogues have a role in a late step WTA synthesis in *B. subtilis* as shown by genetics and by measurements of the WTA content of a triple mutant strain; (ii) the cocrystal structures of *S. pneumoniae* Cps2A containing bound substrate and product analogues that are consistent with the missing enzymatic step; (iii) our demonstration that TagT and Cps2A can both produce a monophosphorylated lipid from a diphosphorylated substrate in a Mg^{2+} -dependent fashion; (iv) the observation of a Mg^{2+} ion in a location indicating a role in phosphorolysis; and (v) mutation of key amino acids in lipid and Mg^{2+} binding blocks the *in vitro* phosphorolysis and abolishes function *in vivo*. Consistent with this conclusion, LCP genes found in AP-containing Gram-positive bacteria are always located in close proximity, often in the same operon, to genes required for AP synthesis. Furthermore, mutants lacking members of the LCP family have been isolated in a number of different bacteria, and in many cases, these mutants have altered cell wall structures, autolysin activities, lower antibiotic resistance levels and/or reduced virulence (Hübscher *et al*, 2008, 2009). Because of these pleiotropic effects, some members of the family were suggested to have a regulatory role in cell wall growth, and the situation is complicated further by the presence in *S. aureus* of a transcriptional activator named LytR that shows no sequence similarity to LCP proteins (Brunskill and Bayles, 1996). The short cytoplasmic regions of a Cps2A homologue from *Streptococcus iniae* were shown to bind DNA *in vitro* (Hanson *et al*, 2011). In addition, early work on the *B. subtilis* *lytR* gene suggested a role in transcriptional regulation because disruption of *lytR* resulted in increased transcription of both *lytR* and the adjacent *lytABC* operon (Lazarevic *et al*, 1992). However, our work does not support a regulatory role for the LCP proteins. Presumably, many of the published phenotypes of LCP mutant strains of the various species studied are indirect consequences of major cell wall alterations and the resulting stress response due to failure of the attachment of APs to PG.

We were unable to detect free APs in the supernatant of the triple *tagTUV* depletion strain (RAD, unpublished data). Indeed, several previous reports have documented a reduction of cell wall-attached APs in LCP mutants of various bacterial species (Massidda *et al*, 1996; Cieslewicz *et al*, 2001; Morona *et al*, 2006; Hübscher *et al*, 2009). It seems possible that feedback regulation blocks the synthesis and/or export of APs when the attachment to PG is impaired.

Implications for the phosphotransferase mechanism catalysed by LCP proteins

The LCP proteins presumably attach teichoic acid, teichuronic acid, or acidic capsular polysaccharide to cell wall PG via phosphodiester linkage to *N*-acetylmuramic acid. Our data on the three *B. subtilis* homologues indicate that the enzymes are not absolutely specific for the type of AP, teichoic acid, or teichuronic acid they attach, because a major phenotype does not emerge until all three genes are inactivated. This is

supported by our crystal structures of Cps2A in complex with reaction product and substrate mimics, which indicate that the enzymes interact strongly with the pyrophosphoryl-lipid component of the AP precursor. There is space in the crystal structure beyond the terminal phosphate of the bound *opr-PP* to accommodate the oligosaccharide moiety that links the lipid to the acidic/AP component. Important interactions with the sugar that is linked directly to the pyrophosphate might occur from the conserved arginine pair, Arg362 and Arg244 of Cps2A. The first two carbohydrate residues of the AP, called the linkage unit, would project into a cleft on the surface of Δ TM-Cps2A (Figure 6E) and the rest of the polymeric, main chain component of the AP precursor is unlikely to interact with the protein. Thus, teichoic/teichuronic acid-attaching enzymes would only recognize the linkage unit of the precursor and not the main chain explaining the lack of specificity of the *Bacillus* enzymes. The PG polymer could be accommodated in the groove above β -strands 18 and 19 (Figure 6E). While we have been able to demonstrate the phosphorolysis of the scissile bond in a substrate mimic, we have not been able to detect phosphotransfer to a PG acceptor. Perhaps, the carbohydrate linkage unit participates in catalysis, by interacting with conserved arginines. If interactions between the linkage unit of AP and protein atoms of the LCP enzyme are crucial to the biochemical reaction, it would help to explain why phosphotransfer efficiency is poor with just a phosphate anion as a leaving group; after all, the reaction trajectories for phosphate monoesters differ fundamentally from those of phosphate di- and tri-esters (Cleland and Hengge, 2006).

The crystal structures provide several clues towards the likely catalytic mechanism. First, Mg^{2+} appears to be an essential catalytic requirement. The divalent cation and the presence of several conserved arginines (e.g. Arg244, Arg267, and Arg362; Figure 5) may contribute to catalysis by stabilizing the transition state by the neutralization of the developing negative charge. Second, there is no obvious amino acid to act as a base in the activation of the incoming nucleophile. There are two acidic groups close to the scissile bond, but both these, Asp234 and Asp246, coordinate the bound magnesium. The coordination of the cation by these two acidic amino acids, and the interaction of phosphate groups with neighbouring positively charged amino acids, is reminiscent of the reaction scheme catalysed by acyl carrier protein (ACP) synthase (Bunkoczi *et al*, 2007). That residues Glu181 and Lys185 participate in general acid and base catalysis in ACP synthase were confirmed by mutagenesis. Possible functional equivalents in Cps2A are Asp234, demonstrated herein to be a crucial amino acid for LCP function, and Arg362, one of several conserved positively charged residues coordinating the pyrophosphate. An alternative candidate to act as the catalytic base is the conserved Asp268, but this is about 5 Å away from the likely position of the incoming nucleophile. Significant conformational changes would be needed to locate this amino acid in a position to extract a proton from phosphotransfer acceptor. Other phosphotransfer reactions do, however, take place without the need for a classic general acid/base pair (Cleland and Hengge, 2006). Finally, the metal ion itself may activate the nucleophilic hydroxyl moiety (Lassila *et al*, 2011) as the phosphoacceptor approaches the active site in order to provide the necessary rate enhancement expected of an enzyme.

We have been unable to visualize the interactions of PG and AP building blocks with the protein, primarily because of the unavailability of some complex substrate components, but also because the active site of Cps2A is already blocked with lipids. It remains to be shown how the lipid is loaded and released from the protein *in vivo*. Perhaps, the phospholipids of the cell membrane can passively empty the enzyme active site without recourse to another protein. Alternatively, the product of the transfer reaction, a monophosphorylated polyisoprenoid lipid might be actively removed from the active site, by an as yet unidentified protein, to be 'flipped' back across the membrane to its inner face so that the phospholipid can be recharged by the cytoplasmic Tag pathway for the enzyme to perform another round of phosphotransfer.

Fungi and plants also covalently attach polysaccharides and other polymers to the primary cell wall but the enzymes catalysing the attachment are not yet known, nor are the mechanisms of spatiotemporal regulation of these processes understood (McNeil *et al*, 1984; Latge, 2007). Thus, our work could stimulate further studies of the assembly of complex cell walls in other organisms.

Global role of the MreB cytoskeleton in coordinating synthesis and assembly of a range of different cell wall polymers

Synthesis of both PG and WTA is concentrated at a central growth zone in *S. pneumoniae* (Tomasz *et al*, 1975), a bacterium that does not contain an MreB cytoskeleton. The discovery of the TagTUV proteins, other WTA proteins, and PG synthetic enzymes in association with the three MreB proteins reinforces the idea that in rod-shaped bacteria, the cytoskeleton orchestrates the activity of not just the PG-synthesizing machinery but also a range of other cell wall components. In support of this idea, several enzymes of both the PG and WTA synthetic machineries have a pattern of localization that appears similar to that of the MreB proteins (Scheffers *et al*, 2004; Claessen *et al*, 2008; Formstone *et al*, 2008) (Supplementary Figure S2). Close association of the PG- and WTA-synthesizing machineries with the MreB proteins could provide a means of coordinating the insertion of WTA with the deposition of PG. Therefore, these results lend strong support to the emerging picture that the MreB cytoskeleton plays a pivotal role in organizing and coordinating the different elements of cell envelope morphogenesis in bacteria.

A novel antibiotic target

Several authors have speculated about the WTA biosynthetic pathway as a potential antibiotic target (for a review, see Swoboda *et al*, 2010). Indeed, Swoboda *et al* (2009) have used a cell-based screen to identify a specific inhibitor of the WTA pathway in *S. aureus*, which was shown to work on the ABC transporter orthologous to TarG in *B. subtilis*. Since the active site of the putative transferase lies outside the cytoplasmic membrane, potential antibiotics targeted on this enzyme would have the advantage of not needing to cross the cytoplasmic membrane. This external location is one of the reasons why β -lactams, targeting the penicillin-binding proteins, have been so successful. Our discovery of the transferase responsible for transfer of nascent WTA to the PG, together with structural information

on the substrate and clues to the likely catalytic mechanism, provides a highly attractive new target for drug discovery programmes.

Materials and methods

Complete details of all the materials and methods used are provided in the Supplementary data.

Bacterial strains, plasmids, and primers

Strains and plasmids used in this study are listed in Supplementary Table SIV and primers are listed in Supplementary Table SV. The construction of plasmids is described in the Supplementary data.

Media

Cultures of *Bacillus* strains were grown in Luria-Bertani (LB) or Difco antibiotic medium 3 (PAB) liquid at 30 or 37 °C, or on LB or nutrient agar (Oxoid) plates at 37 °C, with appropriate supplements when necessary.

Microscopic imaging

For fluorescence microscopy, cells from overnight cultures were diluted into fresh PAB medium and grown to mid-exponential phase at 30 °C. For live cell imaging, cells were mounted on microscope slides covered with a thin film of 1.2% agarose in water, essentially as described previously (Glaser *et al*, 1997). Images were acquired with a Sony Cool-Snap HQ cooled CCD camera attached to a Zeiss Axiovert 200M microscope. The images were acquired and analysed with METAMORPH version 6 (Molecular Devices).

Purification of protein complexes

Proteins that copurified from liquid cultures of *B. subtilis* expressing MreB-, Mbl-, or MreBH-His were purified as described previously (Ishikawa *et al*, 2006; Kawai *et al*, 2009b). After separation by electrophoresis, polyacrylamide gels were divided into slices, and the presence of proteins confirmed by peptide-mass fingerprinting as described previously (Kuwana *et al*, 2002). As a negative control, cells expressing His-tagged Noc, a chromosome-associated negative regulator of cell division (Wu *et al*, 2009), were also analysed. Several proteins were simultaneously identified in both MreB and Noc complexes (e.g. TufA and ribosomal proteins), suggesting that they represent a non-specific background.

Extraction of cell wall material

B. subtilis cell cultures were grown in LB to late exponential phase ($OD_{600} = 0.7$). The cells were harvested, suspended in water, and then immediately added to 10 volumes of 4% SDS at 100 °C. After 20 min incubation, the cell wall material was washed with water 10 times, to remove the SDS. The resultant crude preparation was then suspended in 0.5% NaCl and broken using a bead beater (Bio101 unit at max power for 30 s). After removing the glass beads, the suspension was sedimented by centrifugation, and then washed with water 10 times. The cell wall material was freeze dried in preweighed tubes and to each tube an appropriate amount of water was added to equalize the concentration in each sample. Organic phosphorous was assayed as described previously (Harwood *et al*, 1990).

Electrophoresis of teichoic acids

Purified cell wall samples were treated with 1 M HCl for 20 min at 65 °C and then neutralized with NaOH. Glycerol was added to each hydrolysed sample to give a final concentration of 5%, along with bromophenol blue (0.001% final concentration). Samples were electrophoresed at 10 V/cm on a 12% polyacrylamide gel made in $1.2 \times$ TBE buffer. The gel was stained with Alcian Blue, followed by silver staining, as described previously (Wolters *et al*, 1990).

Protein expression and purification for crystallization

Recombinant Δ TM-Cps2A (residues 98–481, with a C-terminal hexahistidine tag) was produced in the methionine auxotroph B834 (DE3) strain of *E. coli*. Cells were grown in LB media to an OD_{600} of 0.4 and induced with 1 mM IPTG for 18 h at 16 °C for native protein, or in minimal medium supplemented with selenomethionine for the seleno-labelled protein. Recombinant protein was purified as described in the Supplementary data.

Crystallization, structure determination, and refinement of Δ TM-Cps2A

Crystals of Δ TM-Cps2A were obtained by sparse-matrix screening from purified protein at 28 mg/ml. Diffraction data, collected at the Diamond Light Source were processed with MOSFLM (Leslie, 2006) and scaled with SCALA (Evans, 2006). The structure was solved using SHELX (Sheldrick, 2010) within HKL2MAP (Pape and Schneider, 2004); initial model building was performed using ARP/wARP (Langer *et al*, 2008). Refinement proceeded with cycles of automated refinement in PHENIX (Adams *et al*, 2010) and manual rebuilding in COOT (Emsley *et al*, 2010). Details of the diffraction data and refined models are reported in Supplementary Table SII.

Mass spectroscopy of Δ TM-Cps2A

Mass spectroscopy of Δ TM-Cps2A was performed on a Waters LCT Premier XE, using electrospray ionization in both native and denatured states.

Accession codes

Atomic co-ordinates and native structure factors have been deposited at the PDB for the wild type complexes with dpr-*P*, opr-*PP* and opr-*PP*/magnesium and for the R267A mutant of Δ TM-Cps2A in complex with opr-*PP* and manganese with respective accession codes 2xxp, 3tfl, 3tep and 3tel.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (<http://www.embojournal.org>).

References

- Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**: 213–221
- Atilano ML, Pereira PM, Yates J, Reed P, Veiga H, Pinho MG, Filipe SR (2010) Teichoic acids are temporal and spatial regulators of peptidoglycan cross-linking in *Staphylococcus aureus*. *Proc Natl Acad Sci USA* **107**: 18991–18996
- Bhavsar AP, Erdman LK, Schertzer JW, Brown ED (2004) Teichoic acid is an essential polymer in *Bacillus subtilis* that is functionally distinct from teichuronic acid. *J Bacteriol* **186**: 7865–7873
- Bouhss A, Trunkfield AE, Bugg TD, Mengin-Lecreulx D (2008) The biosynthesis of peptidoglycan lipid-linked intermediates. *FEMS Microbiol Rev* **32**: 208–233
- Brown S, Zhang YH, Walker S (2008) A revised pathway proposed for *Staphylococcus aureus* wall teichoic acid biosynthesis based on *in vitro* reconstitution of the intracellular steps. *Chem Biol* **15**: 12–21
- Brunskill EW, Bayles KW (1996) Identification and molecular characterization of a putative regulatory locus that affects auto-lysis in *Staphylococcus aureus*. *J Bacteriol* **178**: 611–618
- Bunkoczi G, Pasta S, Joshi A, Wu X, Kavanagh KL, Smith S, Oppermann U (2007) Mechanism and substrate recognition of human holo ACP synthase. *Chem Biol* **14**: 1243–1253
- Carballido-Lopez R, Formstone A, Li Y, Ehrlich SD, Noirot P, Errington J (2006) Actin homolog MreBH governs cell morphogenesis by localization of the cell wall hydrolase LytE. *Dev Cell* **11**: 399–409
- Cieslewicz MJ, Kasper DL, Wang Y, Wessels MR (2001) Functional analysis in type Ia group B *Streptococcus* of a cluster of genes involved in extracellular polysaccharide production by diverse species of streptococci. *J Biol Chem* **276**: 139–146
- Claessen D, Emmins R, Hamoen LW, Daniel RA, Errington J, Edwards DH (2008) Control of the cell elongation-division cycle by shuttling of PBP1 protein in *Bacillus subtilis*. *Mol Microbiol* **68**: 1029–1046
- Cleland WW, Hengge AC (2006) Enzymatic mechanisms of phosphate and sulfate transfer. *Chem Rev* **106**: 3252–3278
- D'Elia MA, Henderson JA, Beveridge TJ, Heinrichs DE, Brown ED (2009) The N-acetylmannosamine transferase catalyzes the first

Acknowledgements

This work was supported by a grant from the UK Biotechnology and Biological Sciences Research Council to JE, RAD, WV, and RJL. We thank the various group members for helpful discussions, and particularly Leendert Hamoen, Susan Firbank, and Romain Mercier for critical reading of the manuscript and Alice Eberhardt for help with cloning of *cps2A*. We also thank Hiroki Yamamoto, Junichi Sekiguchi, Mark Leaver, and Takuya Morimoto for the gift of strains, and Didier Blanot for gift of undecaprenyl phosphate standard, and Dan Herschlag for communicating a preprint of his review before publication. We also acknowledge the Diamond Light Source for access to its beamlines, the beamline staff and Arnaud Baslé for support during diffraction data collection. We also thank Adam Dowle, Andrew Leech, and Berni Strongitharm of the York University Technology Facility for mass spectroscopic analyses. The authors extend their condolences to the family of Berni Strongitharm, who sadly passed away in May 2011.

Author contributions: RE, SI, MK, NH, NKB, and CNH performed the experiments. YK, JM-W, and RMC designed and performed the experiments, analysed the data, and contributed to the drafting of the paper. NO contributed to the design of the experiments. RJL, WV, RAD, and JE contributed to the planning of the experiments, interpretation of data, and drafting of the paper.

Conflict of interest

The authors declare that they have no conflict of interest.

- committed step of teichoic acid assembly in *Bacillus subtilis* and *Staphylococcus aureus*. *J Bacteriol* **191**: 4030–4034
- D'Elia MA, Millar KE, Beveridge TJ, Brown ED (2006a) Wall teichoic acid polymers are dispensable for cell viability in *Bacillus subtilis*. *J Bacteriol* **188**: 8313–8316
- D'Elia MA, Pereira MP, Chung YS, Zhao W, Chau A, Kenney TJ, Sulavik MC, Black TA, Brown ED (2006b) Lesions in teichoic acid biosynthesis in *Staphylococcus aureus* lead to a lethal gain of function in the otherwise dispensable pathway. *J Bacteriol* **188**: 4183–4189
- Daniel RA, Errington J (2003) Control of cell morphogenesis in bacteria: two distinct ways to make a rod-shaped cell. *Cell* **113**: 767–776
- Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**: 486–501
- Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**: 72–82
- Figge RM, Divakaruni AV, Gober JW (2004) MreB, the cell shape-determining bacterial actin homologue, co-ordinates cell wall morphogenesis in *Caulobacter crescentus*. *Mol Microbiol* **51**: 1321–1332
- Formstone A, Carballido-Lopez R, Noirot P, Errington J, Scheffers DJ (2008) Localization and interactions of teichoic acid synthetic enzymes in *Bacillus subtilis*. *J Bacteriol* **190**: 1812–1821
- Formstone A, Errington J (2005) A magnesium-dependent *mreB* null mutant: implications for the role of *mreB* in *Bacillus subtilis*. *Mol Microbiol* **55**: 1646–1657
- Foster SJ, Popham DL (2002) Structure and synthesis of cell wall, spore cortex, teichoic acids, S-layers, and capsules. In *Bacillus subtilis and Its Closest Relatives: From Genes to Cells*, Sonenshein L, Losick R, Hoch JA (eds) pp 21–41. Washington, DC: American Society for Microbiology
- Glaser P, Sharpe ME, Raether B, Perego M, Ohlsen K, Errington J (1997) Dynamic, mitotic-like behavior of a bacterial protein required for accurate chromosome partitioning. *Genes Dev* **11**: 1160–1168
- Graumann PL (2009) Dynamics of bacterial cytoskeletal elements. *Cell Motil Cytoskeleton* **66**: 909–914
- Hanson BR, Lowe BA, Neely MN (2011) Membrane topology and DNA-binding ability of the *Streptococcal* CpsA protein. *J Bacteriol* **193**: 411–420

- Harwood CR, Coxon RD, Hancock IC (1990) The *Bacillus* cell envelope and secretion. In *Molecular Biology Methods for Bacillus*, Harwood CR, Cutting SM (eds) pp 328–369. New York City: John Wiley & Sons Ltd
- Henriques AO, Glaser P, Piggot PJ, Moran Jr CP (1998) Control of cell shape and elongation by the *rodA* gene in *Bacillus subtilis*. *Mol Microbiol* **28**: 235–247
- Hu B, Yang G, Zhao W, Zhang Y, Zhao J (2007) MreB is important for cell shape but not for chromosome segregation of the filamentous cyanobacterium *Anabaena* sp. PCC 7120. *Mol Microbiol* **63**: 1640–1652
- Hübscher J, Luthy L, Berger-Bachi B, Stutzmann Meier P (2008) Phylogenetic distribution and membrane topology of the LytR-CpsA-Psr protein family. *BMC Genomics* **9**: 617
- Hübscher J, McCallum N, Sifri CD, Majcherczyk PA, Entenza JM, Heusser R, Berger-Bachi B, Stutzmann Meier P (2009) MsrR contributes to cell surface characteristics and virulence in *Staphylococcus aureus*. *FEMS Microbiol Lett* **295**: 251–260
- Ishikawa S, Kawai Y, Hiramatsu K, Kuwano M, Ogasawara N (2006) A new FtsZ-interacting protein, YlmF, complements the activity of FtsA during progression of cell division in *Bacillus subtilis*. *Mol Microbiol* **60**: 1364–1380
- Jones LJ, Carballido-Lopez R, Errington J (2001) Control of cell shape in bacteria: helical, actin-like filaments in *Bacillus subtilis*. *Cell* **104**: 913–922
- Kawai Y, Asai K, Errington J (2009a) Partial functional redundancy of MreB isoforms, MreB, Mbl and MreBH, in cell morphogenesis of *Bacillus subtilis*. *Mol Microbiol* **73**: 719–731
- Kawai Y, Daniel RA, Errington J (2009b) Regulation of cell wall morphogenesis in *Bacillus subtilis* by recruitment of PBP1 to the MreB helix. *Mol Microbiol* **71**: 1131–1144
- Kruse T, Bork-Jensen J, Gerdes K (2005) The morphogenetic MreBCD proteins of *Escherichia coli* form an essential membrane-bound complex. *Mol Microbiol* **55**: 78–89
- Kuwana R, Kasahara Y, Fujibayashi M, Takamatsu H, Ogasawara N, Watabe K (2002) Proteomics characterization of novel spore proteins of *Bacillus subtilis*. *Microbiology* **148**: 3971–3982
- Langer G, Cohen SX, Lamzin VS, Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* **3**: 1171–1179
- Lassila JK, Zalatan JG, Herschlag DL (2011) Biological phosphoryl-transfer reactions: understanding mechanism and catalysis. *Annu Rev Biochem* **80**: 669–702
- Latge JP (2007) The cell wall: a carbohydrate armour for the fungal cell. *Mol Microbiol* **66**: 279–290
- Lazarevic V, Karamata D (1995) The *tagGH* operon of *Bacillus subtilis* 168 encodes a two-component ABC transporter involved in the metabolism of two wall teichoic acids. *Mol Microbiol* **16**: 345–355
- Lazarevic V, Margot P, Soldo B, Karamata D (1992) Sequencing and analysis of the *Bacillus subtilis* *lytRABC* divergon: a regulatory unit encompassing the structural genes of the N-acetylmuramoyl-L-alanine amidase and its modifier. *J Gen Microbiol* **138**: 1949–1961
- Leaver M, Errington J (2005) Roles for MreC and MreD proteins in helical growth of the cylindrical cell wall in *Bacillus subtilis*. *Mol Microbiol* **57**: 1196–1209
- Leslie AG (2006) The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr* **62**: 48–57
- Massidda O, Kariyama R, Daneo-Moore L, Shockman GD (1996) Evidence that the PBP 5 synthesis repressor (*psr*) of *Enterococcus hirae* is also involved in the regulation of cell wall composition and other cell wall-related properties. *J Bacteriol* **178**: 5272–5278
- McNeil M, Darvill AG, Fry SC, Albersheim P (1984) Structure and function of the primary cell walls of plants. *Annu Rev Biochem* **53**: 625–663
- Morona JK, Morona R, Paton JC (2006) Attachment of capsular polysaccharide to the cell wall of *Streptococcus pneumoniae* type 2 is required for invasive disease. *Proc Natl Acad Sci USA* **103**: 8505–8510
- Neuhaus FC, Baddiley J (2003) A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in gram-positive bacteria. *Microbiol Mol Biol Rev* **67**: 686–723
- Pape T, Schneider TR (2004) HKL2MAP: a graphical user interface for phasing with SHELX programs. *J Appl Cryst* **37**: 843–844
- Scheffers DJ, Jones LJ, Errington J (2004) Several distinct localization patterns for penicillin-binding proteins in *Bacillus subtilis*. *Mol Microbiol* **51**: 749–764
- Schirner K, Errington J (2009) The cell wall regulator σI specifically suppresses the lethal phenotype of *mbl* mutants in *Bacillus subtilis*. *J Bacteriol* **191**: 1404–1413
- Schirner K, Stone LK, Walker S (2011) ABC transporters required for export of wall teichoic acids do not discriminate between different main chain polymers. *ACS Chem Biol* **6**: 407–412
- Sheldrick GM (2010) Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr D Biol Crystallogr* **66**: 479–485
- Slovak PM, Wadhams GH, Armitage JP (2005) Localization of MreB in *Rhodobacter sphaeroides* under conditions causing changes in cell shape and membrane structure. *J Bacteriol* **187**: 54–64
- Soldo B, Lazarevic V, Karamata D (2002) *tagO* is involved in the synthesis of all anionic cell-wall polymers in *Bacillus subtilis* 168. *Microbiology* **148**: 2079–2087
- Strahl H, Hamoen LW (2010) Membrane potential is important for bacterial cell division. *Proc Natl Acad Sci USA* **107**: 12281–12286
- Swoboda JG, Campbell J, Meredith TC, Walker S (2010) Wall teichoic acid function, biosynthesis, and inhibition. *ChemBiochem* **11**: 35–45
- Swoboda JG, Meredith TC, Campbell J, Brown S, Suzuki T, Bollenbach T, Malhowski AJ, Kishony R, Gilmore MS, Walker S (2009) Discovery of a small molecule that blocks wall teichoic acid biosynthesis in *Staphylococcus aureus*. *ACS Chem Biol* **4**: 875–883
- Tomasz A, McDonnell M, Westphal M, Zanati E (1975) Coordinated incorporation of nascent peptidoglycan and teichoic acid into pneumococcal cell walls and conservation of peptidoglycan during growth. *J Biol Chem* **250**: 337–341
- Vats P, Shih YL, Rothfield L (2009) Assembly of the MreB-associated cytoskeletal ring of *Escherichia coli*. *Mol Microbiol* **72**: 170–182
- Wei Y, Havasy T, McPherson DC, Popham DL (2003) Rod shape determination by the *Bacillus subtilis* class B penicillin-binding proteins encoded by *pbpA* and *pbpH*. *J Bacteriol* **185**: 4717–4726
- Weidenmaier C, Peschel A (2008) Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. *Nat Rev Microbiol* **6**: 276–287
- Wolters PJ, Hildebrandt KM, Dickie JP, Anderson JS (1990) Polymer length of teichuronic acid released from cell walls of *Micrococcus luteus*. *J Bacteriol* **172**: 5154–5159
- Wu LJ, Ishikawa S, Kawai Y, Oshima T, Ogasawara N, Errington J (2009) Noc protein binds to specific DNA sequences to coordinate cell division with chromosome segregation. *EMBO J* **28**: 1940–1952
- Yokoyama K, Mizuguchi H, Araki Y, Kaya S, Ito E (1989) Biosynthesis of linkage units for teichoic acids in gram-positive bacteria: distribution of related enzymes and their specificities for UDP-sugars and lipid-linked intermediates. *J Bacteriol* **171**: 940–946
- Young M, Mauel C, Margot P, Karamata D (1989) Pseudo-allelic relationship between non-homologous genes concerned with biosynthesis of polyglycerol phosphate and polyribitol phosphate teichoic acids in *Bacillus subtilis* strains 168 and W23. *Mol Microbiol* **3**: 1805–1812